

Ostruzioni topologiche alla coerenza

Una Teoria Combinatoria della Vaghezza

Nota preliminare

Questo saggio sviluppa un quadro matematico per lo studio della vaghezza linguistica mediante strumenti della topologia combinatoria, dell'algebra omologica e della teoria spettrale dei grafi. I punti di partenza — la teoria dei complessi simpliciali, la coomologia, l'omologia persistente, la bandwidth dei grafi, le passeggiate casuali — appartengono al dominio della matematica consolidata, e per ciascuno si forniscono i riferimenti bibliografici precisi. La costruzione che collega questi strumenti al paradosso di Sorite, la definizione del numero di risoluzione ϵ^* , la nozione di sorites barcode e la corrispondenza tra ϵ^* e la struttura dei cammini monotoni nel grafo di tolleranza sono, per quanto è noto all'autore, proposte in questa forma sistematica per la prima volta, sebbene singoli aspetti parziali abbiano precedenti nella letteratura che vengono indicati esplicitamente.

I teoremi enunciati sono accompagnati da dimostrazioni complete. Laddove la dimostrazione dipende da risultati noti, questo è indicato con i riferimenti appropriati. Le proposizioni etichettate come congetture rappresentano affermazioni plausibili per le quali non si dispone di una dimostrazione completa; in tutti i casi viene fornita un'analisi delle strategie di dimostrazione note e dei punti in cui esse si arrestano. L'obiettivo non è risolvere il paradosso di Sorite — operazione che richiederebbe scelte filosofiche che la matematica non può compiere autonomamente — ma descriverne la struttura con la massima precisione formale, rendendo così esplicito quale sia il problema reale e in che termini possa essere quantificato.

Il saggio si sviluppa su più livelli di astrazione, e l'ordine di presentazione è stato scelto per rendere accessibile l'esposizione a lettori con background diverso. I Capitoli 1–4 richiedono soltanto la teoria dei grafi elementare e l'algebra lineare di base; si definiscono gli spazi di tolleranza, i grafi di tolleranza e si enuncia e dimostra il Teorema principale di impossibilità. I Capitoli 5–8 richiedono familiarità con la topologia algebrica elementare — omologia simpliciale, sequenze esatte di Mayer–Vietoris — e sviluppano la teoria del numero di risoluzione ϵ^* e delle ostruzioni omologiche ai predicati ciclicamente vaghi. I Capitoli 9–12 utilizzano l'omologia persistente, la teoria spettrale degli operatori e alcuni elementi di geometria riemanniana discreta; qui si costruisce il sorites barcode e si studia la teoria spettrale della vaghezza. I Capitoli 13–15 trattano applicazioni, aspetti computazionali e connessioni con la linguistica formale, la teoria delle categorie e la semantica compositazionale.

Capitolo 0. Prerequisiti di Topologia Algebrica

0.1 Complessi simpliciali e la loro omologia: dalla combinatoria all'algebra

Questo capitolo riepiloga i fondamenti di topologia algebrica utilizzati nel saggio, con l'obiettivo di rendere il testo autocontenuto per lettori con background in algebra lineare e teoria dei grafi ma senza una formazione specifica in topologia. Le dimostrazioni dei risultati fondamentali sono indicate con i riferimenti bibliografici essenziali; le costruzioni rilevanti per le applicazioni semantiche sono esposte in dettaglio.

Definizione 0.1 (Complesso simpliciale astratto). Un complesso simpliciale astratto su un insieme finito V di vertici è una famiglia K di sottoinsiemi finiti di V che soddisfa: (i) $\{v\} \in K$ per ogni $v \in V$ (ogni singoletto è un simpleso); (ii) se $\sigma \in K$ e $\tau \subseteq \sigma$, allora $\tau \in K$ (chiusura per sottoinsiemi). Un elemento $\sigma \in K$ di cardinalità $k+1$ è un k -simpleso; la dimensione di K è $\dim(K) = \max\{k : \exists \sigma \in K \text{ con } |\sigma| = k+1\}$.

L'operatore di bordo è il cuore della costruzione omologica. Per ogni $k \geq 1$ e ogni k -simpleso $\sigma = [v_0, v_1, \dots, v_k]$ (dove si sceglie un ordinamento dei vertici), il bordo è:

$$\partial_k(\sigma) = \sum_{i=0}^k (-1)^i \cdot [v_0, \dots, \hat{v}_i, \dots, v_k]$$

dove $\hat{v}_i$ indica l'omissione del vertice v_i . L'estensione per linearità su tutte le combinazioni lineari intere di k -simplessi (il gruppo $C_k(K; \mathbb{Z})$ delle k -catene) dà un omomorfismo $\partial_k: C_k(K; \mathbb{Z}) \rightarrow C_{k-1}(K; \mathbb{Z})$.

Lemma 0.1 ($\partial\partial = 0$). Per ogni $k \geq 2$ vale l'identità $\partial_{k-1} \circ \partial_k = 0$, ovvero il bordo del bordo è zero.

Dimostrazione. Si calcola $\partial_{k-1}(\partial_k([v_0, \dots, v_k]))$. Sostituendo la definizione si ottiene una somma doppia: $\sum_i \sum_{j < i} (-1)^{i+j} [v_0, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_k] + \sum_i \sum_{j > i} (-1)^{i+j-1} [v_0, \dots, \hat{v}_i, \dots, \hat{v}_j, \dots, v_k]$. Ogni termine $[v_0, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_k]$ con $i > j$ compare due volte con segni opposti: una dalla coppia (i, j) con contributo $(-1)^{i+j}$ e una dalla coppia (j, i) con contributo $-(-1)^{i+j}$. Le cancellazioni danno il risultato. Si veda Munkres (1984), Lemma 5.1. $\square$

Poiché il bordo del bordo è zero, si ha la catena di complessi:

$$\dots \rightarrow C_{k+1}(K; \mathbb{Z}) \xrightarrow{\partial_{k+1}} C_k(K; \mathbb{Z}) \xrightarrow{\partial_k} C_{k-1}(K; \mathbb{Z}) \rightarrow \dots$$

Il k -esimo gruppo di omologia è per definizione:

$$H_k(K; \mathbb{Z}) = \ker(\partial_k) / \operatorname{im}(\partial_{k+1}) = Z_k(K) / B_k(K)$$

Gli elementi di $Z_k = \ker(\partial_k)$ sono i k -cicli (catene con bordo zero); gli elementi di $B_k = \operatorname{im}(\partial_{k+1})$ sono i k -bordi (catene che sono bordo di qualcosa). Un ciclo che non è bordo di nessuna catena di dimensione superiore rappresenta un "buco" k -dimensionale del complesso. Il rango di $H_k(K; \mathbb{Z})$ come gruppo abeliano libero è il k -esimo numero di Betti β_k .

0.2 Calcolo esplicito per il ciclo C_4

Si calcola l'omologia di $S(C_4, \sim_1)$, il complesso di Sorite del ciclo $C_4 = (\{1,2,3,4\}, \sim)$ con $i \sim j$ se $|i-j| = 1 \pmod{4}$. Il complesso ha 4 vertici e 4 archi: $\{1,2\}, \{2,3\}, \{3,4\}, \{4,1\}$. Non ci sono 2-simplessi. Le catene di grado 0 sono $C_0 = \mathbb{Z}^4$ con base $[1],[2],[3],[4]$. Le catene di grado 1 sono $C_1 = \mathbb{Z}^4$ con base $[12],[23],[34],[41]$ (dove $[ij]$ indica l'arco orientato da i a j).

La matrice di $\partial_1: C_1 \rightarrow C_0$ rispetto alle basi ordinate è:

$$\partial_1 = \begin{vmatrix} -1 & 0 & 0 & 1 \end{vmatrix} \text{ riga } [1]$$

$$\begin{vmatrix} 1 & -1 & 0 & 0 \end{vmatrix} \text{ riga } [2]$$

$$\begin{vmatrix} 0 & 1 & -1 & 0 \end{vmatrix} \text{ riga } [3]$$

$$\begin{vmatrix} 0 & 0 & 1 & -1 \end{vmatrix} \text{ riga } [4]$$

Il nucleo di ∂_1 si calcola risolvendo il sistema lineare. La somma di tutte le righe è zero, dunque il vettore $v = (1,1,1,1)^T$ è nel nucleo. Per calcolare la dimensione, si osserva che la matrice ha rango 3 (le prime tre colonne sono linearmente indipendenti, ma la quarta è la somma delle altre con segni alternati). Dunque $\ker(\partial_1)$ ha dimensione 1, generato da v . Questo corrisponde al ciclo fondamentale $[12]+[23]+[34]+[41]$.

Il gruppo $H_1(C_4; \mathbb{Z}) = \ker(\partial_1)/\operatorname{im}(\partial_2) = \mathbb{Z}/\{0\} \cong \mathbb{Z}$. Il generatore è il ciclo $[12]+[23]+[34]+[41]$. Poiché $\partial_0 = 0$ (non ci sono (-1) -simplessi), $H_0 = C_0/\operatorname{im}(\partial_1) = \mathbb{Z}^4/\operatorname{im}(\partial_1)$. L'immagine di ∂_1 ha rango 3, dunque $H_0 \cong \mathbb{Z}$. Il complesso è connesso (una sola componente connessa) e ha esattamente un ciclo indipendente, coerentemente con il numero cicломatico $\mu = 4-4+1 = 1$.

0.3 Il teorema dei coefficienti universali

Teorema 0.1 (Coefficienti Universali). Per ogni complesso simpliciale K e ogni gruppo abeliano G , esiste una sequenza esatta corta scissa:

$$0 \rightarrow H_k(K; \mathbb{Z}) \otimes G \rightarrow H_k(K; G) \rightarrow \text{Tor}(H_{k-1}(K; \mathbb{Z}), G) \rightarrow 0$$

La scissione non è naturale ma la sequenza lo è. In particolare: (i) se $G = \mathbb{Q}$ o $G = \mathbb{R}$ o $G = \mathbb{Z}/p\mathbb{Z}$ con p primo, $\text{Tor}(H_{k-1}, G)$ si calcola dalla torsione di H_{k-1} ; (ii) se tutti i $H_k(K; \mathbb{Z})$ sono privi di torsione (come per i complessi di flag dei grafi bipartiti), allora $H_k(K; G) \cong H_k(K; \mathbb{Z}) \otimes G$ per ogni G .

Per i complessi di Sorite $S(X, \sim)$ con X finito totalmente ordinato e $\sim$ riflessiva simmetrica, il teorema dei coefficienti universali implica che $H_1(S(X, \sim); \mathbb{R}) \cong \mathbb{R}^{\text{rk}(H_1(S(X, \sim); \mathbb{Z}))}$ ogni volta che $H_1(K; \mathbb{Z})$ è libero (il che vale per tutti i complessi 1-dimensionali, cioè tutti i complessi di Sorite di predicati unidimensionali). Questo sarà usato nel Capitolo 8 per la teoria spettrale.

0.4 Sequenze esatte corte e lunghe

Uno strumento essenziale per il calcolo dell'omologia di spazi composti è la sequenza esatta lunga di Mayer-Vietoris. Siano A e B sottocomplessi di K con $A \cup B = K$. La sequenza esatta lunga è:

$$\dots \rightarrow H_{k+1}(A \cup B) \xrightarrow{\delta} H_k(A \cap B) \xrightarrow{\{i_a, i_b\}} H_k(A) \oplus H_k(B) \xrightarrow{\{j_a - j_b\}}$$

$$H_k(A \cup B) \rightarrow H_{k-1}(A \cap B) \rightarrow \dots$$

dove i_a, i_b sono le inclusioni di $A \cap B$ in A e B rispettivamente, j_a, j_b sono le inclusioni di A e B in $A \cup B$, e δ è il morfismo di connessione (snake lemma). L'esattezza della sequenza — cioè il fatto che l'immagine di ogni mappa coincide col nucleo della mappa successiva — segue dal lemma del serpente applicato alla sequenza corta di complessi di catene:

$$0 \rightarrow C(A \cap B) \rightarrow C(A) \oplus C(B) \rightarrow C(A \cup B) \rightarrow 0$$

Si veda Hatcher (2002), Sezione 2.2, per la dimostrazione completa. L'applicazione ai complessi di Sorite è sistematica nel Capitolo 6.

0.5 Il rango del gruppo di omologia H_1 per grafi: la formula del numero ciclomattico

Proposizione 0.1 (Formula del numero ciclomático per complessi 1-dimensionali). Sia $G = (V, E)$ un grafo connesso finito. Il complesso simpliciale K con 0-simplessi V e 1-simplessi E è 1-dimensionale (non ci sono 2-simplessi). Il rango di $H_1(K; \mathbb{Z})$ come gruppo abeliano libero è:

$$\beta_1 = \text{rk}(H_1(K; \mathbb{Z})) = |E| - |V| + 1 = \mu(G)$$

dove $\mu(G)$ è il numero ciclomático del grafo G .

Dimostrazione. Dalla sequenza esatta $0 \rightarrow H_1 \rightarrow C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} H_0 \rightarrow 0$ (con $H_0 \cong \mathbb{Z}$ per la connessione) e dalla formula dei ranghi per sequenze esatte corte: $\text{rk}(H_1) = \text{rk}(\ker \partial_1) = \text{rk}(C_1) - \text{rk}(\text{im } \partial_1) = |E| - (|V| - 1) = |E| - |V| + 1$. Il termine $\text{rk}(\text{im } \partial_1) = |V| - 1$ segue dal teorema rango-nullità: ∂_1 ha nucleo di dimensione β_1 e immagine di dimensione $|E| - \beta_1$; d'altra parte $\text{im}(\partial_1) \subseteq C_0 = \mathbb{Z}^{|V|}$ ha rango al più $|V|$, ed è esattamente $|V| - 1$ perché la somma di tutte le colonne di ∂_1 è zero (la somma dei bordi di ogni albero di copertura è zero in C_0). Questo è consistente con la formula di Euler $\chi = |V| - |E| = \beta_0 - \beta_1 = 1 - \mu$. $\square$

La Proposizione 0.1 è lo strumento principale per il calcolo di H_1 nei Capitoli 6 e 7. In particolare, il numero di ostruzioni omologiche indipendenti al paradosso ciclico è esattamente $\mu(G(X, \sim))$: il numero di cicli indipendenti nel grafo di tolleranza.

Capitolo 1. Il Paradosso di Sorite come Problema Matematico

1.1 La formulazione canonica e la sua struttura logica

Il paradosso di Sorite — dal greco σωρός, mucchio — è documentato per la prima volta nella tradizione degli Scettici e dei Megarici ed è attribuito a Eubulide di Mileto nel IV secolo a.C., lo stesso logico cui si devono altri celebri paradossi dell'antichità, tra cui quello del Mentitore. La formulazione che è diventata lo standard di riferimento nella filosofia contemporanea si articola in 2 premesse e 1 conclusione:

P1: 1 grano di sabbia non forma un mucchio. P2: Se n grani non formano un mucchio, nemmeno $n + 1$ lo formano. C: Nessun numero di grani forma un mucchio.

La deduzione da P1 e P2 alla conclusione C è un'applicazione diretta del principio di induzione matematica: si parte dal caso base P1 ($n = 1$) e si applica iterativamente P2 per ottenere che 2, 3, 4, ... grani non formano un mucchio, e così per ogni n finito. La struttura formale della deduzione è inattaccabile: si tratta di un'applicazione corretta di un principio logico valido in qualunque sistema deduttivo sufficientemente ricco da includere l'induzione sui naturali.

Eppure l'enunciato "10.000 grani formano un mucchio" è vero per qualunque uso ordinario del termine mucchio. Anzi, lo è già per numeri molto più piccoli: 500 grani disposti in modo appropriato formano chiaramente un mucchio nel senso in cui questa parola viene normalmente impiegata. Il ragionamento è logicamente valido — la deduzione per induzione è correttamente costruita — ma conduce a una conclusione universalmente riconosciuta come falsa, e quindi almeno uno dei 2 assunti deve essere falso. Il paradosso ha dunque la struttura formale di una *reductio ad absurdum*, ma con il peculiare rovesciamento che è la conclusione apparentemente ovvia — l'esistenza di mucchi — a essere paradossale, non la premessa iniziale.

La difficoltà è che entrambe le premesse sembrano accettabili. P1 descrive un caso paradigmatico di non-mucchio: nessun parlante di lingua italiana identificherebbe 1 grano di sabbia come un mucchio in qualunque contesto ordinario. P2 cattura l'intuizione che la differenza tra n e $n + 1$ grani sia semanticamente irrilevante per il predicato "mucchio": se la classificazione come mucchio dipendesse dalla differenza di un singolo grano, dovrebbe esistere un intero n^* tale che n^* grani non formano un mucchio ma $n^* + 1$ grani sì. Questa soglia sembra arbitraria in un senso molto preciso: non c'è nulla nel significato ordinario di "mucchio" che stabilisca dove tale confine debba trovarsi, e la sua negazione è precisamente P2.

La struttura logica del paradosso è quella di un'antinomia semantica, ovvero l'incoerenza di un insieme di assunti che individualmente sembrano tutti ragionevoli. La congiunzione di P1, P2 e del fatto osservato che esistono mucchi è incoerente in qualunque logica che includa il principio di induzione. In logica classica bivalente, almeno 1 dei 3 elementi deve essere abbandonato: o P1 (ma allora 1 grano forma un mucchio, il che è linguisticamente inaccettabile), o P2 (ma allora esiste un confine netto, che però sembra arbitrario), o si nega l'esistenza dei mucchi (ma questa è la conclusione paradossale che volevamo evitare).

La teoria matematica proposta in questo saggio non risolve questa scelta filosofica, ma la descrive con precisione strutturale: mostra quale configurazione topologica del dominio del predicato rende la scelta inevitabile e misura quantitativamente il costo di ciascuna delle 3 opzioni. Questo è il senso preciso in cui la matematica contribuisce al problema della vaghezza: non dicendo quale premessa abbandonare, ma rendendo esplicito il prezzo da pagare per ciascuna scelta.

1.2 La pervasività del fenomeno e la sua indipendenza dal predicato specifico

Lo stesso schema si ripresenta per una classe vastissima di predicati del linguaggio naturale, e questa pervasività è uno dei motivi per cui il fenomeno merita un trattamento matematico sistematico piuttosto che una casistica filosofica predicato per predicato.

Un individuo alto 150 cm è chiaramente non alto secondo tutti gli usi ordinari dell'italiano in un contesto di statura maschile adulta; uno alto 200 cm è chiaramente alto. Nessuna differenza di 1 mm tra 2 stature sembra sufficiente a stabilire una classificazione diversa: se una persona di 183 cm è "alta", perché quella di 182,9 cm non lo sarebbe? Il predicato "alto" soggiace alla stessa struttura paradossale del predicato "mucchio".

Lo stesso vale per "calvo"rispetto al numero di capelli (0 capelli è chiaramente calvo, 100.000 capelli non lo è, ma dove è il confine?), "giovane"rispetto all'età in anni, "rosso"rispetto alla lunghezza d'onda in nanometri (dove comincia l'arancio?), "ricco"rispetto al patrimonio in euro, "veloce"rispetto alla velocità in km/h, "caldo"rispetto alla temperatura in gradi Celsius. Per ciascuno di questi predicati, il paradosso di Sorite si formula in modo identico cambiando soltanto i termini specifici.

Un aspetto decisivo che distingue la vaghezza da altre forme di indeterminatezza semantica è che il fenomeno non dipende da ignoranza fattuale. Nel caso del predicato "mucchio", si sa esattamente quanti grani ci sono, si sa che differiscono di 1, ma non si sa cosa dire del predicato. Se l'incertezza fosse fattuale, basterebbe contare meglio. Invece l'incertezza è semantica: il predicato stesso non determina dove passa il confine. Questa caratteristica distingue la vaghezza dall'ambiguità (dove esistono 2 o più significati distinti che il contesto non disambigua, come "banco"che può significare istituto di credito o sedile) e dall'indeterminatezza referenziale (dove il termine si riferisce a più individui senza che sia chiaro quale, come "il professore di matematica" in un contesto con più professori di matematica). La vaghezza è una proprietà strutturale del predicato in relazione al suo dominio, non una mancanza di informazione sul mondo.

Una classe importante di predicati vaghi ha struttura multidimensionale, e questa è la classe per cui la teoria topologica sviluppata nel seguito è strettamente necessaria. "È in forma fisicamente"dipende simultaneamente da più parametri: la percentuale di grasso corporeo, la capacità aerobica, la forza muscolare, la flessibilità. Non solo i parametri sono molteplici, ma non si ordinano naturalmente su una scala unica: una persona potrebbe avere eccellenti capacità aerobiche e scarsa flessibilità, oppure buona forza muscolare e mediocre resistenza cardiovascolare. "È intelligente"dipende da capacità cognitive diverse — ragionamento logico, memoria, intelligenza verbale, pensiero spaziale, intelligenza emotiva — che la ricerca psicologica ha dimostrato non essere riducibili a un'unica dimensione nonostante i tentativi del coefficiente g. "È bello"dipende da innumerevoli tratti percettivi e culturali.

Per questi predicati multidimensionali, la struttura della vaghezza è più ricca di quella dei predicati unidimensionali in un senso matematico preciso: il complesso simpliciale associato può avere ostruzioni omologiche in gradi più alti, e il numero di risoluzione dipende dalla geometria di uno spazio di alta dimensione piuttosto che da un semplice percorso su una retta.

Un'ulteriore classe di predicati vaghi presenta struttura ciclica. "È di sinistra" su uno spettro politico che si chiude su sé stesso in alcune descrizioni — dove l'estrema sinistra e l'estrema destra condividono alcune posizioni anti-establishment — è un esempio. "È rosso" su uno spettro cromatico chiuso è un altro: la percezione cromatica forma un cerchio (l'hue circle della teoria del colore), non una retta, e il predicato "è caldo come colore" si applica tanto al rosso quanto all'arancio quanto al viola, che percettivamente si avvicina di nuovo al rosso. Per questi predicati, il paradosso di Sorite ha una forma diversa dalla versione lineare: non si tratta di distinguere 2 estremi su una retta, ma di trovare un assegnamento coerente in presenza di cicli di indistinguibilità. La teoria omologica sviluppata nei Capitoli 5–7 si occupa sistematicamente di questo caso, mostrando come la presenza di cicli nel grafo di tolleranza produca ostruzioni omologiche non banali che impediscono assegnamenti coerenti non costanti.

1.3 Le risposte filosofiche principali

Le risposte filosofiche al paradosso di Sorite si articolano in 3 grandi tradizioni. Il loro esame è necessario non soltanto per contestualizzare il quadro matematico, ma perché ciascuna tradizione corrisponde a una scelta precisa sulla struttura topologica del dominio del predicato, e il linguaggio degli spazi di tolleranza permette di rendere queste scelte esplicite e confrontabili.

L'epistemicismo, difeso in modo sistematico e rigoroso da Timothy Williamson in *Vagueness* (1994) e da Roy Sorensen in *Blindspots* (1988), sostiene che i predicati vaghi hanno confini netti. Esiste un intero n^* tale che n^* grani non formano un mucchio ma $n^* + 1$ grani sì; esiste un'altezza esatta h^* tale che chiunque misuri meno di h^* non è alto ma chiunque misuri h^* o più lo è. La nostra incapacità di determinare n^* o h^* non è un difetto della realtà ma della nostra conoscenza: siamo epistemicamente limitati nell'identificare il confine esatto.

Williamson propone una semantica del significato basata sull'uso complessivo della comunità linguistica nel tempo: l'insieme di tutti gli usi del termine da parte di tutti i parlanti nel tempo determina un confine preciso, anche se nessun parlante individuale è in grado di conoscerlo. La vaghezza è dunque un fenomeno epistemologico, non semantico o ontologico. Il vantaggio è la conservazione integrale della logica classica bivalente e dell'induzione: $P2$ è semplicemente falsa, anche se non possiamo sapere per quale n . Il problema principale è spiegare perché il confine esista dove è — in base a quale proprietà della realtà o dell'uso linguistico — e perché sia sistematicamente inaccessibile alla conoscenza umana. Williamson risponde con la sua semantica dell'anti-luminosità: la nostra conoscenza è sensibile alle condizioni in modo che ci rende incapaci di conoscere casi limite, non per un limite strumentale ma per ragioni strutturali legate alla natura della conoscenza. Sorensen invece fa appello all'opacità cognitiva come caratteristica generale della mente, non specifica alla semantica.

Nel quadro matematico di questo saggio, l'epistemicismo corrisponde all'affermazione che il grafo di tolleranza "vero" del predicato — quello determinato dall'uso completo della comunità linguistica — è disconnesso. Il numero di risoluzione $\epsilon^*(\sim_vero)$ è 0, e la struttura di tolleranza osservata

empiricamente ($\sim$ _osservata) è soltanto un'approssimazione rumorosa di quella vera. La distanza bottleneck tra il sorites barcode del predicato osservato e quello di un predicato epistemico (con confine netto) misura il "costo epistemico" in termini strutturali.

La teoria della supervalutazione, sviluppata sistematicamente da Kit Fine in *Vagueness, Truth and Logic* (1975) e con importanti contributi precedenti di Mehlberg (1958) e van Fraassen (1966), adotta una prospettiva diversa. Un predicato vago come "mucchio" ammette molteplici precisazioni: ciascuna fissa un confine netto (dice: il confine è a n^* grani), ma la scelta tra le precisazioni ammissibili è arbitraria entro certi limiti, nel senso che la semantica del predicato non determina quale precisazione adottare. Un enunciato è supervero se è vero in tutte le precisazioni ammissibili, superfalso se è falso in tutte, e privo di valore di verità determinato negli altri casi.

Il vantaggio della supervalutazione è duplice: si preserva la logica classica a livello delle singole precisazioni, e si cattura l'intuizione che predicati vaghi abbiano una zona di indeterminatezza genuina. Un problema noto (Fara, 2000; Williamson, 1994) è che l'enunciato "Esiste un n tale che n grani non formano un mucchio ma $n + 1$ grani sì" risulta supervero — vero in ogni precisazione, poiché ogni precisazione ammissibile fissa un confine netto — anche se nessuna istanza specifica " n grani non formano un mucchio ma $n + 1$ sì" per un determinato n è supervera. Questo crea tensioni con la logica del primo ordine: l'esistenziale supervero ha istanze tutte non-vere (né vere né false), il che viola il principio classico che un enunciato esistenziale vero implica l'esistenza di un'istanza vera.

Nel quadro del presente saggio, lo spazio delle precisazioni ammissibili corrisponde all'insieme dei tagli del grafo di tolleranza che disconnettono $\min(X)$ da $\max(X)$. Ogni precisazione è un taglio: la rimozione di 1 o più archi del grafo fino a rendere il predicato nitido. Il sorites barcode descrive la struttura di questo spazio: le barre con lunga persistenza in PD_0 corrispondono a tagli robusti (precisazioni stabili che rimangono valide per un ampio intervallo di perturbazioni della metrica), quelle con breve persistenza a tagli fragili.

Le logiche a gradi di verità, nella tradizione di Łukasiewicz (1920), Zadeh (1965) con la logica fuzzy, e Edgington (1997) con la probabilistic vagueness, assegnano a ogni enunciato un valore nell'intervallo reale $[0, 1]$. Le connettive logiche vengono ridefinite per preservare la struttura graduata: la negazione di un enunciato con valore v ha valore $1 - v$, la congiunzione di due enunciati con valori v_1 e v_2 ha valore $\min(v_1, v_2)$ nella logica di Łukasiewicz (o $v_1 \cdot v_2$ in quella di prodotto), la disgiunzione ha valore $\max(v_1, v_2)$.

Queste teorie catturano l'intuizione che la vaghezza sia continua: "è mucchio" assume valore 0 per 1 grano, valori intermedi per quantità intermedie, valore 1 per molti grani. Il problema principale è la sensibilità alla scelta della funzione di assegnazione: perché la funzione lineare piuttosto che quella sigmoide, o quella esponenziale? La teoria non fornisce una risposta intrinseca a questa domanda. Un secondo problema è la cosiddetta obiezione della "super-vaghezza": anche la funzione di membership è soggetta a vaghezza (per quale valore esatto la funzione di membership smette di essere 1?), generando una regressione all'infinito.

Il numero di risoluzione ε^* offre una risposta parziale a entrambi i problemi. Come dimostrato nel Capitolo 5 (Teorema 3 e Proposizione 25), la struttura di tolleranza del predicato determina una funzione di membership ottimale — quella che minimizza ε tra tutte le funzioni ε -coerenti monotone con le condizioni agli estremi — e per tolleranze uniformi questa funzione ottimale è esattamente la funzione lineare di Łukasiewicz. Il numero ε^* quantifica il grado di vaghezza, e la sua derivazione dalla struttura topologica del predicato lo ancora a qualcosa di semanticamente motivato piuttosto che arbitrario.

Il quadro matematico proposto è ortogonale a tutte e 3 le tradizioni: descrive la struttura combinatoria e topologica intrinseca a qualunque predicato vago, indipendentemente dalla semantica prescelta. Qualunque risposta al paradosso venga adottata — epistemica, supervaluzionista o graduata — deve fare i conti con la topologia del complesso associato al predicato. La topologia non risolve il paradosso, ma ne delimita esattamente lo spazio delle soluzioni.

1.4 La domanda matematica

La domanda centrale è: data la struttura di indistinguibilità di un predicato vago — formalizzata come relazione di tolleranza su un insieme finito di istanze — quale struttura combinatoria e topologica ne emerge, e come si misura quantitativamente il grado e il tipo di vaghezza?

La risposta che questo saggio elabora è articolata su 3 livelli. Al primo livello, la struttura rilevante è quella del complesso di Sorite, un complesso simpliciale costruito a partire dalla relazione di tolleranza del predicato. Le proprietà topologiche di questo complesso — connessione, numero di componenti, cicli indipendenti, omologia persistente — misurano in modo preciso e semanticamente neutro il grado e il tipo di vaghezza. Al secondo livello, il numero di risoluzione $\varepsilon^* = 1/L_mono_min$ quantifica in un unico numero reale nell'intervallo $[0, 1]$ il costo minimo di rilassamento della coerenza necessario per ottenere un assegnamento accettabile. Al terzo livello, il sorites barcode — il diagramma di persistenza della filtrazione di Sorite — fornisce una firma topologica del predicato che è stabile rispetto a perturbazioni della metrica (Teorema 12) e che consente di confrontare e classificare predicati diversi.

Il paradosso di Sorite, in questo quadro, non è un fallimento della logica ma una conseguenza necessaria della topologia del complesso associato: è un teorema. Dimostrarne la necessità matematica non lo rende meno paradossale, ma chiarisce dove esattamente risiede la difficoltà e quali siano le alternative precise per affrontarla.

Capitolo 2. Spazi di Tolleranza e Grafi di Tolleranza

2.1 La relazione di tolleranza e la sua giustificazione psicofisica

Definizione 1 (Spazio di tolleranza). Uno spazio di tolleranza è una coppia $(X, \sim)$ dove X è un insieme finito e $\sim$ è una relazione binaria su X che soddisfa le seguenti 2 proprietà:

- Riflessività: $x \sim x$ per ogni x in X .
- Simmetria: se $x \sim y$ allora $y \sim x$.

La relazione $\sim$ non è richiesta essere transitiva. Questo è il punto che distingue uno spazio di tolleranza da uno spazio di equivalenza, ed è precisamente questa assenza di transitività che rende la struttura matematicamente interessante e semanticamente appropriata per modellare la vaghezza.

Per rendere visibile perché la transitività sia assente, vale la pena esaminare in dettaglio l'interpretazione semantica. La relazione $x \sim y$ si legge "x è indistinguibile da y ai fini del predicato P", oppure in modo equivalente "la differenza tra x e y è semanticamente irrilevante per P". Se $\sim$ fosse anche transitiva, sarebbe una relazione di equivalenza, le sue classi di equivalenza sarebbero le "unità atomiche indistinguibili" del predicato, e il problema della vaghezza sarebbe banale: il predicato sarebbe necessariamente costante su ciascuna classe, e la classificazione in Vero/Falso dipenderebbe soltanto a quale classe appartiene l'istanza. Un tale predicato sarebbe "nitido a grana grossa": invece di avere un confine su una singola istanza, ne avrebbe uno su una classe intera. La vaghezza emerge precisamente dalla non-transitività: $x \sim y$ e $y \sim z$ non implicano $x \sim z$, e questo significa che la catena di indistinguibilità può connettere elementi distantissimi nell'ordinamento semantico del predicato. È possibile avere $x_1 \sim x_2 \sim x_3 \sim \dots \sim x_n$ dove x_1 è chiaramente Falso e x_n è chiaramente Vero, ma non esiste alcun passo della catena in cui la distinzione sia linguisticamente giustificata.

La giustificazione empirica della non-transitività viene dalla psicofisica, e in particolare dalla nozione di differenza appena percettibile just noticeable difference (JND), introdotta da Ernst Heinrich Weber nel 1834 e formalizzata da Gustav Theodor Fechner nel 1860. La legge di Weber–Fechner descrive la relazione tra la grandezza fisica di uno stimolo e la sua percezione soggettiva, e in particolare stabilisce che la soglia di discriminazione sensoriale ΔI per uno stimolo di intensità I è proporzionale a I : $\Delta I/I = k$ (costante di Weber), ovvero la JND è una frazione costante dell'intensità corrente.

L'implicazione per la non-transitività è diretta. Siano x_1, x_2, x_3 3 stimoli con $x_1 < x_2 < x_3$ tali che $x_2 - x_1 < \text{JND}(x_1)$ e $x_3 - x_2 < \text{JND}(x_2)$. Allora $x_1 \sim x_2$ (indistinguibili) e $x_2 \sim x_3$ (indistinguibili), ma $x_3 - x_1 = (x_2 - x_1) + (x_3 - x_2)$ può benissimo superare $\text{JND}(x_1)$, nel qual caso x_1 e x_3 sono distinguibili, ossia non vale $x_1 \sim x_3$. La composizione di passi sub-JND produce una differenza super-JND: questa è esattamente la struttura che la teoria degli spazi di tolleranza formalizza.

Il fenomeno è stato documentato sperimentalmente con grande precisione per grandezze fisiche diverse. Per la lunghezza d'onda della luce nel visibile, la JND cromatica è dell'ordine di 1–3 nm nelle regioni di maggiore sensibilità (verde–giallo, intorno a 555 nm) e sale fino a 5–10 nm nelle regioni di minore sensibilità (rosso, viola). Per l'intensità sonora, la JND è di circa 1 dB per suoni di media intensità. Per la statura umana, la JND visiva è dell'ordine di 5–10 mm quando le persone sono affiancate. Per il peso percepito a mano, la costante di Weber è circa 2%, ovvero la JND per un peso di 500 g è circa 10 g.

Tutte queste grandezze producono strutture di tolleranza non transitive in modo sistematico: 2 stimoli che differiscono di meno di 1 JND sono "indistinguibili" (la probabilità che un soggetto li classifichi correttamente non è significativamente diversa da quella del caso), ma composizioni di passi sub-JND producono differenze sopra-JND. E poiché i predicati linguistici come "alto", "rosso", "caldo" sono calibrati sulla percezione sensoriale, le strutture di tolleranza da essi indotte sono non transitive per la stessa ragione.

Un secondo meccanismo che produce non-transitività è la somiglianza prototipica nella categorizzazione concettuale, studiata da Eleanor Rosch a partire dagli anni '70. Secondo la teoria dei prototipi di Rosch, le categorie concettuali non hanno confini netti ma si strutturano intorno a esempi tipici (prototipi) rispetto ai quali gli esempi periferici sono valutati per grado di somiglianza. La relazione di somiglianza è notoriamente non transitiva: Wittgenstein lo aveva osservato nella sua discussione delle Familienähnlichkeiten (somiglianze di famiglia) nelle Philosophische Untersuchungen (1953). Un gioco da tavolo somiglia a un gioco di carte (entrambi hanno regole, si gioca in compagnia, c'è un vincitore), e un gioco di carte somiglia al solitario (entrambi usano le stesse carte, hanno struttura simile), ma un gioco da tavolo e il solitario hanno meno in comune; la transitività della somiglianza manca in modo sistematico nelle tassonomie concettuali.

Osservazione 1. La chiusura transitiva $\sim^T$ di una relazione di tolleranza $\sim$ è la minima relazione di equivalenza che estende $\sim$. Esplicitamente, $x \sim^T y$ se e solo se esiste una sequenza finita $x = z_0, z_1, \dots, z_k = y$ tale che $z_i \sim z_{i+1}$ per ogni i . Le classi di equivalenza di $\sim^T$ corrispondono esattamente alle componenti connesse del grafo di tolleranza $G(X, \sim)$. La vaghezza emerge nel "divario" tra $\sim$ e $\sim^T$: elementi che appartengono alla stessa classe di equivalenza di $\sim^T$ (e dunque alla stessa componente connessa del grafo) ma che non sono direttamente in relazione di tolleranza tra loro sono "indirettamente indistinguibili" attraverso catene di tolleranza. Il Teorema 1 mostra che questa connessione indiretta è sufficiente a produrre il paradosso di Sorite.

La chiusura transitiva $\sim^T$ è computazionalmente accessibile: si calcola con l'algoritmo di Floyd–Warshall in tempo $O(n^3)$ su n elementi, oppure con una visita BFS/DFS del grafo di tolleranza in tempo $O(n + |E|)$. La struttura di $\sim^T$ è quella di una partizione di X in componenti connesse, e la dimensione e struttura di queste componenti è il primo invariante topologico di interesse.

2.2 Il grafo di tolleranza e le sue proprietà strutturali

Ad ogni spazio di tolleranza $(X, \sim)$ è associato in modo canonico il grafo non orientato $G(X, \sim) = (X, E)$ dove $E = \{(x, y) : x \sim y, x \neq y\}$. I cappi corrispondenti alla riflessività ($x \sim x$) sono omessi per convenzione, poiché non aggiungono informazione strutturale: il grafo è semplice e non orientato.

Il grafo di tolleranza codifica tutta l'informazione strutturale dello spazio di tolleranza che è rilevante per la teoria dell'assegnamento coerente. Le sue proprietà topologiche — connessione, numero di componenti connesse, numero ciclomatico, diametro, struttura bipartita o meno, grado dei vertici — hanno interpretazioni dirette in termini di vaghezza del predicato associato, che ora esplicitiamo.

Il numero di componenti connesse k del grafo $G(X, \sim)$ è il primo invariante. Per la Proposizione 1 che si dimostra di seguito, il numero di assegnamenti coerenti $f: X \rightarrow \{\text{Vero}, \text{Falso}\}$ è esattamente 2^k . Un grafo con una singola componente connessa ($k = 1$) ammette soltanto 2 assegnamenti coerenti: la funzione costantemente Vero e quella costantemente Falso. In termini semantici, un predicato con grafo di tolleranza connesso non può distinguere il suo minimo dal suo massimo in modo coerente — questo è precisamente il paradosso di Sorite. Un grafo con k componenti ammette 2^k assegnamenti coerenti, e tra questi ci sono assegnamenti che assegnano valori diversi a componenti diverse, rendendo possibile la distinzione degli estremi se essi appartengono a componenti diverse.

Il diametro di $G(X, \sim)$, definito come $\text{diam}(G) = \max_{\{x, y \in X\}} d_G(x, y)$ dove d_G è la distanza nel grafo (numero minimo di archi in un cammino), misura quanto "lontani" devono essere 2 elementi perché non siano connessi da nessuna catena di tolleranza di lunghezza limitata. Predicati con grafo di tolleranza di diametro grande hanno catene di indistinguibilità molto lunghe: è necessaria una sequenza di molti passi per connettere il minimo al massimo. Come si vedrà, il diametro del DAG monotono (Lemma 2 del Capitolo 5) è inversamente proporzionale al numero di risoluzione ϵ^* .

Il numero ciclomatico $\mu(G) = |E| - |V| + c$, dove c è il numero di componenti connesse di G , misura il numero di cicli indipendenti nel grafo. Per un grafo connesso ($c = 1$), $\mu = |E| - |V| + 1$. Come si vedrà nei Capitoli 5–7, ogni ciclo indipendente contribuisce un'ostruzione omologica indipendente agli assegnamenti coerenti: la presenza di cicli rende impossibile certe classificazioni binarie non per ragioni di connessione (che riguarda H_0) ma per ragioni omologiche di ordine superiore ($H_1, H_2, \dots$).

Il grado $\deg(x)$ di un vertice $x \in X$ è il numero di elementi y con $x \sim y, y \neq x$. In termini semantici, $\deg(x)$ conta quanti altri elementi del dominio sono "indistinguibili" da x . Un elemento con grado alto è "immerso" in molte relazioni di tolleranza: è difficile distinguerlo dai suoi vicini. Un elemento con grado basso è relativamente isolato: è distinguibile da quasi tutti gli altri elementi. Per predicati con struttura di tolleranza uniforme (come "alto" rispetto alla statura), tutti i vertici hanno lo stesso grado (il grafo è regolare).

Proposizione 1 (Numero di assegnamenti coerenti). Il numero di assegnamenti coerenti $f: X \rightarrow \{\text{Vero}, \text{Falso}\}$ — dove per "coerente" si intende che $f(x) = f(y)$ ogni volta che $x \sim y$ — è esattamente 2^k , dove k è il numero di componenti connesse di $G(X, \sim)$.

Dimostrazione. La condizione di coerenza $f(x) = f(y)$ per $x \sim y$ richiede che f sia costante su ogni componente connessa: se x e y sono nella stessa componente, esiste un cammino $x = z_0, z_1, \dots, z_m = y$ in $G(X, \sim)$, e per coerenza $f(z_0) = f(z_1) = \dots = f(z_m)$. Viceversa, se x e y sono in componenti diverse, nessun cammino li collega nel grafo, e la coerenza non impone nessun vincolo tra $f(x)$ e $f(y)$. Le k componenti connesse $C_1, C_2, \dots, C_k$ sono disgiunte e la loro unione è X . La funzione f è dunque completamente determinata dalla scelta di un valore in $\{\text{Vero}, \text{Falso}\}$ per ciascuna delle k componenti, e queste scelte sono indipendenti l'una dall'altra. Ci sono quindi esattamente 2^k assegnamenti coerenti possibili. $\square$

Proposizione 2 (Caratterizzazione per le componenti). Sia $G(X, \sim)$ il grafo di tolleranza con componenti connesse $C_1, C_2, \dots, C_k$. La restrizione di ogni assegnamento coerente f a ciascuna componente C_i è la funzione costante uguale al valore $f(C_i) \in \{\text{Vero}, \text{Falso}\}$ assegnato all'intera componente. Ogni scelta indipendente di $f(C_i)$ per $i = 1, \dots, k$ produce un assegnamento coerente distinto, e tutti i 2^k assegnamenti coerenti si ottengono in questo modo.

Dimostrazione. La costanza su ciascuna componente segue direttamente dalla dimostrazione della Proposizione 1. La biezione tra scelte di valori per le componenti e assegnamenti coerenti è stabilita dalla stessa argomentazione. $\square$

La Proposizione 2 mostra che la struttura interna delle componenti connesse è irrilevante per la classificazione degli assegnamenti coerenti bivalenti: due componenti con struttura interna diversa (una catena, un completo, un ciclo) danno ugualmente contributo di 1 grado di libertà. Questo cambia quando si considerano assegnamenti a valori in $[0, 1]$: la struttura interna diventa rilevante perché determina il numero di risoluzione ε^* , come dimostrato nel Capitolo 5.

La teoria degli assegnamenti bivalenti può essere estesa a valori più generali in modo naturale. Se $V = \{v_1, v_2, \dots, v_m\}$ è un insieme finito di valori di verità, il numero di assegnamenti coerenti $f: X \rightarrow V$ è m^k , dove k è il numero di componenti. Per $V = [0, 1]$ continuo, il problema degli assegnamenti coerenti esatti (0-coerenti) si riduce ai casi $m = 2$ (soltanto le funzioni costanti per componente), ma il problema degli assegnamenti ε -coerenti per $\varepsilon > 0$ ha una ricchezza strutturale molto maggiore, che è il cuore della teoria del Capitolo 5.

2.3 Tolleranze metriche e parametriche

Un caso particolarmente importante e frequente di spazio di tolleranza è quello in cui la relazione $\sim$ è definita a partire da una metrica o pseudometrica su X .

Definizione 2 (Tolleranza metrica). Sia (X, d) uno spazio metrico finito e sia $r > 0$ un parametro reale. La tolleranza di raggio r è la relazione $\sim_r$ definita da $x \sim_r y$ se e solo se $d(x, y) \leq r$. Per ogni $r \geq 0$, $(X, \sim_r)$ è uno spazio di tolleranza: la riflessività è garantita da $d(x, x) = 0 \leq r$ per ogni $r \geq 0$; la simmetria dalla simmetria della metrica.

La tolleranza metrica produce per $r = 0$ la relazione di identità ($x \sim_0 y$ sse $x = y$, poiché $d(x, y) = 0$ sse $x = y$ per una metrica), e per $r \geq \text{diam}(X) = \max_{\{x, y\}} d(x, y)$ la relazione universale ($x \sim_r y$ per ogni x, y). Le tolleranze metriche parametrizzano dunque una famiglia di spazi di tolleranza indexed dal parametro $r \in [0, \infty)$, con la proprietà di monotonia: $r \leq s$ implica $\sim_r \subseteq \sim_s$ nel senso che $x \sim_r y$ implica $x \sim_s y$. Al crescere di r , il grafo di tolleranza acquista archi (non ne perde) e tende progressivamente verso il grafo completo K_n . Questa famiglia è la filtrazione di Sorite introdotta formalmente nel Capitolo 9.

La non-transitività delle tolleranze metriche è garantita dalla geometria: se $d(x, y) \leq r$ e $d(y, z) \leq r$, la disuguaglianza triangolare implica soltanto $d(x, z) \leq 2r$, non $d(x, z) \leq r$. Per spazi metrici in cui la disuguaglianza triangolare è stretta (ad esempio, spazi in cui i punti sono in posizione "generica"), ci sarà tipicamente x e z a distanza $2r$ (o vicino a $2r$) con y a distanza r da entrambi, così che $x \sim_r y$ e $y \sim_r z$ ma non $x \sim_r z$. La non-transitività emerge quindi naturalmente dalla geometria dello spazio metrico, non come ipotesi artificiale imposta dall'esterno.

Definizione 3 (Tolleranza di larghezza uniforme). Se $X = \{x_1 < x_2 < \dots < x_n\}$ è un insieme totalmente ordinato con n elementi e k è un intero positivo con $1 \leq k \leq n - 1$, la tolleranza di larghezza uniforme k è la relazione: $x_i \sim x_j$ se e solo se $|i - j| \leq k$.

Questa definizione modella il caso in cui la distinguibilità dipende soltanto dalla posizione ordinale nell'insieme, non da alcuna metrica continuamente variabile. È il modello più semplice e regolare di spazio di tolleranza unidimensionale, e serve da caso di riferimento per quasi tutti i calcoli espliciti del saggio.

Per $k = 1$, la tolleranza di larghezza uniforme produce il grafo catena $1 - 2 - 3 - \dots - n$: ogni elemento è tollerante soltanto con i suoi 2 vicini immediati (con l'eccezione degli elementi agli estremi, che hanno 1 solo vicino). Questo è il modello della JND con soglia 1 posizione ordinale. Per $k = 2$, ogni elemento è tollerante con i 2 vicini a destra e i 2 a sinistra, e il grafo diventa più "denso". Per $k = n - 1$, ogni elemento è tollerante con ogni altro elemento, producendo il grafo completo K_n .

La maggior parte dei risultati "espliciti" di questo saggio — calcolo di ε^* , formula per $\text{rk}(H_1)$, analisi del sorites barcode — vengono prima dimostrati per tolleranze uniformi e poi estesi a tolleranze arbitrarie, seguendo la strategia consolidata in matematica di partire dal caso più regolare e poi generalizzare.

Definizione 4 (Bandwidth costretta rispetto a un ordinamento). Per uno spazio di tolleranza $(X, \sim, \leq)$ con $X = \{x_1 < x_2 < \dots < x_n\}$ totalmente ordinato, la bandwidth costretta rispetto all'ordinamento $\leq$ è:

$$bw_{\leq}(\sim) = \max \{|i - j| : x_i \sim x_j, i \neq j\}$$

La bandwidth costretta misura il massimo "salto ordinale" tra 2 elementi che sono in relazione di tolleranza. È un indicatore della "non-località" della relazione: una bandwidth piccola indica che ogni elemento è tollerante soltanto con i suoi vicini immediati nell'ordinamento, mentre una bandwidth grande indica la presenza di relazioni di tolleranza tra elementi lontani nell'ordinamento.

Per tolleranze di larghezza uniforme k , si ha banalmente $bw_{\leq} = k$. Per tolleranze arbitrarie, $bw_{\leq}$ può variare da 1 (soltanto i vicini immediati sono tolleranti, come nel grafo catena puro) fino a $n - 1$ (gli elementi agli 2 estremi sono tolleranti, come nella Proposizione 7 del Capitolo 5).

La bandwidth costretta fornisce un upper bound su $\varepsilon^*(\sim)$ — come dimostrato nella Proposizione 5 del Capitolo 5 — ma non è in generale uguale a $\varepsilon^*(\sim)$. La discrepanza tra le 2 quantità, e la struttura dei predicati per cui essa è grande, è uno dei temi che emergono nell'analisi del Capitolo 5 e che rivelano come la struttura globale dei cammini (catturata da ε^*) sia un invariante genuinamente diverso dalla struttura locale delle relazioni di tolleranza (catturata da $bw_{\leq}$).

2.4 Strutture di tolleranza su spazi di caratteristiche continue

Nella pratica della linguistica computazionale e della semantica formale, i predicati vaghi non operano su insiemi discreti finiti ma su spazi di caratteristiche continue. Vale la pena discutere come la teoria degli spazi di tolleranza si estenda a questo contesto, indicando anche i punti in cui emergono difficoltà tecniche.

Sia $X = [0, 1]$ l'intervallo reale unitario, interpretato come il continuo delle altezze maschili normalizzate tra il minimo antropologico e il massimo. La relazione di tolleranza $x \sim_r y$ sse $|x - y| \leq r$ per $r > 0$ è ben definita (riflessiva e simmetrica), non transitiva, e produce per ogni $r > 0$ un grafo di tolleranza che è il grafo di vicinanza su $[0, 1]$ con "larghezza di banda" r . In questo caso continuo, il grafo di tolleranza è un oggetto continuo (non finito), e la teoria algebrica applicabile è quella della topologia algebrica per spazi compatti.

La differenza principale rispetto al caso finito è che per X continuo la filtrazione di Sorite a ogni raggio $r > 0$ produce un complesso di Sorite che è già connesso (il grafo di vicinanza su $[0, 1]$ con larghezza $r > 0$ è connesso): il numero di risoluzione $\varepsilon^* = 0$ per qualunque $r > 0$, e il paradosso di Sorite è "immediato" nel senso che non c'è nessun raggio di tolleranza positivo che consenta un assegnamento coerente non costante. Questo rende il caso continuo meno interessante dal punto di vista della filtrazione e del barcode: il sorites threshold r^* è 0.

La teoria interessante per X continuo emerge in due direzioni. La prima è quella dei campioni finiti: si sceglie un sottoinsieme finito $X_n \subset [0, 1]$ di n punti (un campione di stature misurate, per esempio) e si studia la filtrazione di Sorite su X_n . Al crescere di n , il sorites barcode di X_n converge al sorites barcode di $[0, 1]$ nel senso della distanza bottleneck, con un tasso di convergenza che dipende dalla distribuzione dei punti del campione. Questa è la direzione della Direzione 1 nei Futuri Sviluppi del Capitolo 15. La seconda direzione è quella dei predicati multidimensionali, dove $X \subset \mathbb{R}^d$ per $d \geq 2$: in questo caso il complesso di Sorite può avere ostruzioni omologiche non banali già per $r = 0^+$ (cicli che persistono per un intervallo di raggi positivo), e il barcode ha struttura non triviale. Questa è la direzione principale dei Capitoli 7–9.

Capitolo 3. Il Complesso di Sorite

3.1 Costruzione e proprietà fondamentali

Il complesso di Sorite è l'oggetto topologico centrale della teoria. La sua costruzione generalizza quella del grafo di tolleranza aggiungendo struttura dimensionale: non soltanto le relazioni tra coppie di elementi, ma anche quelle tra triple, quadruple, e insiemi più grandi.

Definizione 5 (Complesso di Sorite). Dato uno spazio di tolleranza $(X, \sim)$, il complesso di Sorite $S(X, \sim)$ è il complesso simpliciale definito come segue: i vertici (0-simplessi) di $S(X, \sim)$ sono gli elementi di X , e un sottoinsieme finito $\{x_0, x_1, \dots, x_k\} \subseteq X$ è un k -simplessodi $S(X, \sim)$ se e solo se $x_i \sim x_j$ per ogni coppia di indici $i \neq j$ in $\{0, 1, \dots, k\}$, ossia se e solo se ogni coppia di elementi del sottoinsieme è in relazione di tolleranza.

In altri termini, $S(X, \sim)$ è il complesso delle clique (o complesso di flag) del grafo di tolleranza $G(X, \sim)$: ogni sottoinsieme di vertici che forma una clique nel grafo (un sottografo completo) diventa un semplice del complesso. La dimensione del semplice è una in meno della cardinalità dell'insieme: una coppia $\{x, y\}$ è un 1-simplesso (arco), una tripla mutualmente tollerante $\{x, y, z\}$ è un 2-simplesso (triangolo), e così via.

Questa costruzione è nota in topologia combinatoria come complesso di flag o complesso delle clique del grafo, ed è uno degli oggetti più studiati della topologia combinatoria moderna. Una delle sue proprietà distintive, il teorema di bandiera (Gal, 2005), afferma che per i complessi di flag la struttura topologica è completamente determinata dal grafo 1-scheletro: due complessi di flag con lo stesso grafo 1-scheletro hanno lo stesso tipo di omotopia. Questo significa che, per gli spazi di tolleranza, tutta la topologia rilevante è già codificata nel grafo di tolleranza $G(X, \sim)$, e il

complesso di Sorite non aggiunge informazione "di ordine superiore" che non sia già presente nel grafo.

Proprietà fondamentale (Chiusura per sottoinsiemi). Il complesso $S(X, \sim)$ è ben definito come complesso simpliciale, nel senso che soddisfa la proprietà di chiusura per sottoinsiemi: ogni sottoinsieme di un simpleso è un simpleso. Se $\{x_0, \dots, x_k\}$ è mutualmente tollerante (ogni coppia è in relazione $\sim$), allora ogni sottoinsieme $\{x_{i_1}, \dots, x_{i_m}\} \subseteq \{x_0, \dots, x_k\}$ è a sua volta mutualmente tollerante (perché ogni coppia del sottoinsieme è anche una coppia del simpleso originale), ed è dunque un m -simpleso di $S(X, \sim)$.

Dimostrazione. Diretta dalla definizione: se $x_i \sim x_j$ per ogni i, j in $\{0, \dots, k\}$, allora in particolare $x_{i_a} \sim x_{i_b}$ per ogni a, b in $\{1, \dots, m\}$ (poiché $\{i_a\}$ e $\{i_b\}$ sono elementi di $\{0, \dots, k\}$). $\square$

La chiusura per sottoinsiemi è la proprietà definitoria dei complessi simpliciali: garantisce che ogni "faccia" di un simpleso sia ancora un simpleso del complesso, il che è necessario perché la realizzazione geometrica sia un CW-complesso con struttura topologica coerente.

Definizione 6 (Realizzazione geometrica). La realizzazione geometrica $|S(X, \sim)|$ è lo spazio topologico ottenuto come segue. Per ogni k -simpleso $\sigma = \{x_0, \dots, x_k\}$ di $S(X, \sim)$, si considera il simpleso geometrico standard $\Delta^k = \{(t_0, \dots, t_k) \in \mathbb{R}^{k+1} : t_i \geq 0, \sum t_i = 1\}$. Si incollano questi semplici geometrici lungo le facce comuni: due semplici geometrici Δ^k e Δ^m corrispondenti a σ e τ vengono incollati lungo la faccia corrispondente a $\sigma \cap \tau$. La topologia risultante su $|S(X, \sim)|$ è la topologia quoziente.

La realizzazione geometrica trasforma il complesso simpliciale combinatorio (un oggetto puramente combinatorio, una famiglia di sottoinsiemi di X) in uno spazio topologico concreto, su cui si possono definire e calcolare i gruppi di omologia. La topologia di $|S(X, \sim)|$ è il principale invariante di interesse: le sue proprietà — tipo di omotopia, gruppi di omologia, numero di componenti — misurano la struttura della vaghezza in modo geometricamente intuitivo e algebricamente preciso.

Osservazione 2 (Relazione con il complesso di Vietoris–Rips). Quando $\sim = \sim_r$ è una tolleranza metrica di raggio r su uno spazio metrico finito (X, d) , il complesso di Sorite $S(X, \sim_r)$ coincide con il complesso di Vietoris–Rips $VR(X, r)$, definito come il complesso simpliciale le cui clique sono i sottoinsiemi di diametro $\leq r$. Questa coincidenza non è casuale ma è conseguenza diretta delle definizioni: un sottoinsieme $\{x_0, \dots, x_k\}$ è una clique di $G(X, \sim_r)$ (ogni coppia ha distanza $\leq r$) se e solo se il diametro di $\{x_0, \dots, x_k\}$ è $\leq r$ (poiché il diametro è il massimo delle distanze tra coppie).

La coincidenza con $VR(X, r)$ permette di applicare direttamente al complesso di Sorite tutti i risultati della teoria dell'omologia persistentesviluppati per la filtrazione di Vietoris–Rips, inclusi i teoremi di approssimazione topologica di Latschev (2001) — che mostra come $VR(X, r)$ approssimi il tipo di omotopia della varietà riemanniana da cui X è campionato, quando r è

sufficientemente piccolo e il campione è sufficientemente denso — e i risultati di stabilità di Cohen-Steiner, Edelsbrunner e Harer (2007) — che mostra la stabilità del diagramma di persistenza rispetto a perturbazioni della metrica, come formalizzato nel Teorema 12 del Capitolo 9.

Il complesso di Vietoris–Rips ha una storia lunga e articolata. Leopold Vietoris lo introdusse nel 1927 come strumento per calcolare l'omologia di spazi metrici compatti mediante approssimazioni finite; Eliyahu Rips lo reintrodusse in modo indipendente negli anni '80 nel contesto della teoria dei gruppi geometrici. La sua applicazione alla topologia algebrica e all'analisi dei dati è diventata sistematica con il lavoro di Edelsbrunner, Letscher e Zomorodian (2002) sulla topologia computazionale.

Esiste anche il complesso di Čech come alternativa al complesso di Vietoris–Rips. Il complesso di Čech $\check{C}(X, r)$ ha come k -simplessi i sottoinsiemi $\{x_0, \dots, x_k\}$ tali che le palle $B(x_i, r)$ hanno intersezione comune non vuota (non solo a coppie, ma tutte insieme). Il lemma di Nerve (Teorema A di seguito) garantisce che $\check{C}(X, r)$ ha lo stesso tipo di omotopia di $\bigcup_x B(x, r)$ (l'unione delle palle). Il complesso di Vietoris–Rips è computazionalmente più trattabile (la condizione "ogni coppia a distanza $\leq r$ " è verificabile in tempo polinomiale) ma approssima la topologia in modo leggermente meno fedele; in particolare, $VR(X, r) \subseteq \check{C}(X, r) \subseteq VR(X, 2r)$ (inclusioni di complessi).

3.2 Esempi fondamentali, tipi di omotopia e significato semantico

È istruttivo esaminare in dettaglio una serie di esempi che illustrano la gamma di strutture topologiche possibili per il complesso di Sorite.

Esempio 1 (Catena di larghezza 1). Sia $X = \{1, 2, 3, 4, 5\}$ con la relazione $i \sim j$ sse $|i - j| \leq 1$. Il grafo di tolleranza è la catena lineare $1 - 2 - 3 - 4 - 5$. Le clique del grafo sono le coppie $\{i, i+1\}$ per $i = 1, 2, 3, 4$ e i singoletti. Non ci sono triangoli (nessuna tripla è mutualmente adiacente: tra i vertici 1, 2, 3, la coppia $\{1, 3\}$ non è un arco). Il complesso è quindi 1-dimensionale, composto da 4 segmenti incollati lungo i vertici comuni.

La realizzazione geometrica $|S(X, \sim)|$ è uno spazio topologico che ha il tipo di omotopia di un punto: è contraibile. Questo significa $H_0 = \mathbb{Z}$ (1 sola componente connessa), $H_k = 0$ per $k \geq 1$ (nessun ciclo, nessun vuoto in dimensione ≥ 1). Il fatto che la catena sia contraibile ha un'interpretazione semantica precisa: la struttura di tolleranza di larghezza 1 su 5 elementi produce il tipo di vaghezza più semplice possibile, in cui gli elementi sono "maximally connected" senza però creare strutture topologiche non banali. Il predicato è massimamente vago (non si può distinguere il minimo dal massimo), ma la vaghezza è "piana" (nessun ciclo, nessuna struttura di ordine superiore).

Il numero di risoluzione ε^* per questa struttura è $1/4$ (il cammino monotono minimo da 1 a 5 ha lunghezza 4), come si calcolerà esplicitamente nel Capitolo 5.

Esempio 2 (Catena di larghezza 2). Sia $X = \{1, 2, 3, 4, 5\}$ con la relazione $i \sim j$ sse $|i - j| \leq 2$. Le clique includono ora le triple: $\{1, 2, 3\}$, $\{2, 3, 4\}$, $\{3, 4, 5\}$ sono tutte clique di dimensione 3 (triangoli nel grafo). Il complesso ha 3 triangoli (2-simpleksi) incollati lungo i lati.

La realizzazione geometrica è ancora contraibile: $H_0 = \mathbb{Z}$, $H_k = 0$ per $k \geq 1$. Aumentare la larghezza da 1 a 2 non produce nuova topologia: i triangoli "riempiono" i cicli eventualmente creati dagli archi aggiuntivi, e il complesso rimane topologicamente equivalente a un punto. Il numero di risoluzione è ora $\varepsilon^* = 1/2$ (il cammino monotono minimo da 1 a 5 ha lunghezza 2: il cammino $1 \rightarrow 3 \rightarrow 5$).

Questo esempio illustra un principio generale: l'aggiunta di relazioni di tolleranza a maggiore "portata" (archi più lunghi) tende a ridurre il numero di risoluzione ε^* (rendendo più facile distinguere gli estremi in senso graduato) senza necessariamente cambiare la topologia del complesso.

Esempio 3 (Ciclo C_6). Sia $X = \{1, 2, 3, 4, 5, 6\}$ con la distanza circolare $d_{\text{circ}}(i, j) = \min(|i - j|, 6 - |i - j|)$ e la relazione $i \sim_1 j$ sse $d_{\text{circ}}(i, j) \leq 1$. Il grafo è il ciclo C_6 : $1 - 2 - 3 - 4 - 5 - 6 - 1$. Non ci sono triangoli: i vicini di 1 nel grafo sono 2 e 6, che non sono adiacenti tra loro (la distanza circolare tra 2 e 6 è 2, non 1). In C_6 , nessuna tripla è mutualmente adiacente per la stessa ragione.

Il complesso è 1-dimensionale (nessun simplekso di dimensione ≥ 2) e il suo tipo di omotopia è quello di S^1 (il cerchio): $H_0 = \mathbb{Z}$ (1 componente), $H_1 = \mathbb{Z}$ (1 ciclo fondamentale indipendente), $H_k = 0$ per $k \geq 2$. La presenza di $H_1 \cong \mathbb{Z}$ è la prima ostruzione omologica non banale: essa indica che il complesso ha un "buco" 1-dimensionale che non può essere riempito da 2-simpleksi (perché non ce ne sono). Semanticamente, questo corrisponde a un paradosso ciclico: non esiste nessun assegnamento coerente non costante su C_6 , e qualunque tentativo di assegnare valori diversi a elementi diversi viene bloccato dalla struttura del ciclo.

Esempio 4 (Simplekso pieno). Se $\sim$ è la relazione universale (ogni coppia è tollerante: $x \sim y$ per ogni $x, y \in X$), allora $S(X, \sim)$ è il simplekso pieno su X : un unico $(|X| - 1)$ -simplekso. Il simplekso pieno è contraibile: $H_0 = \mathbb{Z}$, $H_k = 0$ per $k \geq 1$. In termini semantici, questo corrisponde a un predicato massimamente vago: qualunque assegnamento coerente deve essere costante su tutto X . Non soltanto è impossibile distinguere il minimo dal massimo, ma è impossibile distinguere qualunque coppia di elementi.

Esempio 5 (Complesso discreto). Se $\sim$ è la relazione di identità ($x \sim y$ sse $x = y$), il complesso è il complesso discreto: soltanto 0-simpleksi, nessun arco. $H_0 = \mathbb{Z}^n$ (dove $n = |X|$, e ogni elemento è la propria componente connessa), $H_k = 0$ per $k \geq 1$. Ogni elemento è distinto da ogni altro, e il numero di assegnamenti coerenti è 2^n (uno per ogni funzione $X \rightarrow \{\text{Vero}, \text{Falso}\}$). Questo corrisponde a un predicato completamente nitido, senza nessuna relazione di indistinguibilità tra elementi diversi.

Esempio 6 (Griglia 3×3). Sia $X = \{(i, j) : 1 \leq i, j \leq 3\}$ con la relazione di adiacenza di Manhattan-1: $(i, j) \sim (i', j')$ sse $|i - i'| + |j - j'| \leq 1$. Il grafo ha 9 vertici e 12 archi (4 archi interni più 8 archi di bordo). Non ci sono triangoli in questo grafo: i 4 vicini ortogonali di (i, j) sono $(i \pm 1, j)$ e $(i, j \pm 1)$, e la distanza di Manhattan tra 2 vicini ortogonali qualsiasi di (i, j) è 2, non 1, quindi non sono adiacenti tra loro.

Il numero ciclomatico è $\mu = |E| - |V| + 1 = 12 - 9 + 1 = 4$. Ci sono 4 cicli quadrati indipendenti, corrispondenti alle 4 facce quadrate della griglia 3×3 . $H_0 = \mathbb{Z}$ (connessa), $H_1 = \mathbb{Z}^4$ (4 cicli indipendenti), $H_k = 0$ per $k \geq 2$. Questo esempio illustra come la vaghezza bidimensionale produca una struttura omologica molto più ricca di quella unidimensionale: 4 ostruzioni omologiche indipendenti invece di 1.

Esempio 7 (Bordo dell'ottaedro). Si consideri il grafo completo K_4 su 4 vertici meno 1 arco: $X = \{1, 2, 3, 4\}$ con tutti gli archi eccetto $\{1, 4\}$. Le clique sono: tutti i singoletti, tutti gli archi eccetto $\{1, 4\}$, e i triangoli $\{1, 2, 3\}$ e $\{2, 3, 4\}$ (che sono clique perché tutti gli archi interni sono presenti). Non c'è il tetraedro pieno perché manca l'arco $\{1, 4\}$. Il complesso ha 2 triangoli incollati lungo il lato $\{2, 3\}$.

La realizzazione geometrica è il "diamante" o "rombo": 2 triangoli incollati lungo 1 lato. Questa è topologicamente S^1 (un anello quadrato). $H_0 = \mathbb{Z}$, $H_1 = \mathbb{Z}$, $H_k = 0$ per $k \geq 2$. L'ostruzione è nuovamente in H_1 , non in H_2 .

Esempio 8 (Bordo del tetraedro come 2-sfera). Si consideri $X = \{1, 2, 3, 4\}$ con la relazione $\sim_{\text{bordo}}$ in cui tutti gli archi sono presenti (il grafo completo K_4), ma i semplici del complesso sono limitati al bordo del 3-simplesso: si includono tutti i k -semplici per $k \leq 2$ ma non il 3-simplesso $\{1, 2, 3, 4\}$ stesso. Questa struttura non corrisponde al complesso di flag del grafo completo (che includerebbe automaticamente il 3-simplesso), e quindi non è un complesso di Sorite nel senso della Definizione 5.

Questo esempio illustra un punto sottile: i complessi di Sorite sono sempre complessi di flag, il che significa che ogni insieme i cui elementi sono mutualmente tolleranti è un semplice. Non si può "mettere a mano" quali semplici includere; la struttura è completamente determinata dal grafo. Per ottenere $H_2 \neq 0$ come ostruzione semantica, si avrebbe bisogno di una struttura di tolleranza in cui certe triple sono mutualmente tolleranti ma il 3-simplesso risultante non "riempie" un ciclo 2-dimensionale — ma questo richiederebbe un grafo in cui certi triangoli mancano. Come discusso nel Capitolo 6, le ostruzioni in H_2 richiedono strutture di tolleranza molto particolari.

Esempio 9 (Il predicato "è di media altezza" come struttura bimodale). Consideriamo $X = \{1, \dots, 11\}$ con una struttura di tolleranza non uniforme: $x_i \sim x_j$ se i e j sono entrambi nel segmento $\{1, \dots, 5\}$ e $|i - j| \leq 2$, oppure entrambi nel segmento $\{7, \dots, 11\}$ e $|i - j| \leq 2$, oppure $\{i, j\} = \{5, 7\}$ (connettore tra i 2 blocchi). Il valore $x = 6$ è "di media altezza" e viene escluso dal modello.

Il grafo di tolleranza ha 2 blocchi densi ($\{1, \dots, 5\}$ e $\{7, \dots, 11\}$) connessi da 1 arco $\{5, 7\}$, e il vertice 6 è isolato. Il complesso ha 3 componenti connesse (i 2 blocchi e il vertice 6), ed è topologicamente equivalente a 2 intervalli più un punto. $H_0 = \mathbb{Z}^3$ (3 componenti), $H_1 = \mathbb{Z}^4$ (2 cicli per blocco da ciascun blocco denso). Il numero di risoluzione per distinguere 1 da 11 è $\varepsilon^* = 1/L_{\text{mono_min}}$ dove $L_{\text{mono_min}}$ è il cammino monotono minimo da 1 a 11 attraverso il connettore $\{5, 7\}$.

3.3 Il teorema del nervo e le sue implicazioni

Un risultato fondamentale della topologia combinatoria, il teorema del nervo di Leray (1945), permette di confrontare il tipo di omotopia del complesso di Sorite con quello dello spazio geometrico sottostante, fornendo la base teorica per le applicazioni della TDA all'analisi dei dati.

Teorema A (Teorema del nervo). Sia $U = \{U_\alpha\}_\alpha$ uno ricoprimento aperto di uno spazio topologico Y tale che ogni intersezione finita non vuota $\cap \{\alpha \in F\} U_\alpha$ sia contraibile (o equivalentemente aciclicare) per ogni sottoinsieme finito F . Allora il nervo $N(U)$ — il complesso simpliciale le cui semplici sono le sottofamiglie finite di U con intersezione non vuota — ha lo stesso tipo di omotopia di Y .

Dimostrazione. Si veda Borsuk (1948) o Hatcher (2002), Capitolo 4, per la dimostrazione nel caso aperto. Il caso più generale in cui le intersezioni sono solo acicliche richiede un argomento di successione spettrale; si veda Munkres (1984), Sezione 70. $\square$

Applicazione al complesso di Sorite. Sia (X, d) uno spazio metrico finito immerso in $\mathbb{R}^d$, con la tolleranza metrica $\sim_r$. Per ogni $x \in X$, sia $U_x = B(x, r/2) \subset \mathbb{R}^d$ la palla aperta di raggio $r/2$ centrata in x . Il nervo $N(U)$ delle palle $U = \{U_x\}_{x \in X}$ è il complesso di Čech $\check{C}(X, r/2)$. Per punti in posizione generica in $\mathbb{R}^d$ (soddisfacenti la condizione di general position), ogni intersezione finita di palle è convessa (per il lemma di Helly applicato a convessi in $\mathbb{R}^d$), dunque contraibile. Il teorema del nervo dà allora: $|\check{C}(X, r/2)|$ ha lo stesso tipo di omotopia di $\bigcup_{x \in X} B(x, r/2)$.

Il complesso di Vietoris–Rips $VR(X, r) = S(X, \sim_r)$ contiene $\check{C}(X, r)$ come sottocomplesso e a sua volta è contenuto in $\check{C}(X, 2r)$: in simboli, $\check{C}(X, r) \subseteq VR(X, 2r) \subseteq \check{C}(X, 2r)$. Queste inclusioni garantiscono che il tipo di omotopia di $|VR(X, r)|$ è "vicino" a quello di $\bigcup_x B(x, r/2)$, con un errore di fattore 2 nel raggio. Per applicazioni pratiche (dove r viene variato nella filtrazione), questa approssimazione è accettabile e il complesso di Sorite fornisce un'approssimazione computazionalmente trattabile della topologia dell'unione delle palle.

Teorema B (Latschev, 2001). Sia M una varietà riemanniana compatta, e sia $X \subset M$ un campione ε -denso (ogni punto di M è a distanza $\leq \varepsilon$ da qualche punto di X). Per r sufficientemente piccolo (in relazione al raggio di injectività di M e alla curvatura) e per X sufficientemente denso (ε

sufficientemente piccolo rispetto a r), il complesso di Vietoris–Rips $VR(X, r)$ ha lo stesso tipo di omotopia di M .

Dimostrazione. Latschev (2001). Il risultato più preciso, con quantificazione esplicita di "sufficientemente piccolo", è in Niyogi, Smale e Weinberger (2008). $\square$

Il Teorema di Latschev è la base teorica per l'uso della TDA per ricostruire la topologia di spazi continui da campioni finiti: se i punti sono sufficientemente densi, il complesso di Sorite approssima fedelmente la topologia dello spazio sottostante. Per l'analisi della vaghezza, questo significa che il complesso di Sorite costruito su un campione di istanze del predicato approssima la topologia dello "spazio dei valori" del predicato, e le ostruzioni omologiche del complesso riflettono le ostruzioni topologiche intrinseche del dominio continuo.

3.4 Morfismi tra spazi di tolleranza e funzorialità del complesso di Sorite

La costruzione del complesso di Sorite non è soltanto un'operazione che produce un oggetto topologico da uno spazio di tolleranza: è un funtore tra categorie, nel senso della teoria delle categorie. Questo punto è importante perché la funzorialità garantisce che le proprietà topologiche del complesso di Sorite siano invarianti rispetto alle trasformazioni naturali tra predicati.

Definizione 7 (Morfismo di spazi di tolleranza). Siano $(X, \sim_X)$ e $(Y, \sim_Y)$ due spazi di tolleranza. Una funzione $\varphi: X \rightarrow Y$ è un morfismo di spazi di tolleranza se $x \sim_X x'$ implica $\varphi(x) \sim_Y \varphi(x')$ per ogni $x, x' \in X$. È un isomorfismo se è biiettiva e preserva biunivocamente la relazione di tolleranza: $x \sim_X x'$ se e solo se $\varphi(x) \sim_Y \varphi(x')$.

I morfismi di spazi di tolleranza sono le "trasformazioni che rispettano la struttura di indistinguibilità": se x e x' sono indistinguibili per il predicato P su X , allora le loro immagini $\varphi(x)$ e $\varphi(x')$ sono indistinguibili per il predicato Q su Y . In termini semantici, φ è una "traduzione" del predicato P nel predicato Q che preserva la struttura di vaghezza.

Proposizione 3 (Funzorialità del complesso di Sorite). Ogni morfismo $\varphi: (X, \sim_X) \rightarrow (Y, \sim_Y)$ induce un'applicazione simpliciale $S(\varphi): S(X, \sim_X) \rightarrow S(Y, \sim_Y)$ che manda ogni k -simplexso $\{x_0, \dots, x_k\}$ nel k -simplexso $\{\varphi(x_0), \dots, \varphi(x_k)\}$. A sua volta, $S(\varphi)$ induce omomorfismi $\varphi^*: H_k(S(X, \sim_X)) \rightarrow H_k(S(Y, \sim_Y))$ per ogni $k \geq 0$. Questa costruzione è funtoriale: l'identità va nell'identità e la composizione di morfismi va nella composizione degli omomorfismi.

Dimostrazione. Se $\{x_0, \dots, x_k\}$ è un k -simplexso di $S(X, \sim_X)$, allora $x_i \sim_X x_j$ per ogni $i \neq j$. Poiché φ è un morfismo, $\varphi(x_i) \sim_Y \varphi(x_j)$ per ogni $i \neq j$. Dunque $\{\varphi(x_0), \dots, \varphi(x_k)\}$ è un k -simplexso di $S(Y, \sim_Y)$. L'applicazione $S(\varphi)$ manda simplexsi in simplexsi, è quindi un'applicazione simpliciale. La funzorialità segue dalla funzorialità standard dei gruppi di omologia rispetto alle applicazioni simpliciali. $\square$

Proposizione 4 (Categoria degli spazi di tolleranza). Gli spazi di tolleranza finiti con i morfismi della Definizione 7 formano una categoria, denotata TolSp . Il complesso di Sorite S è un funtore $S: \text{TolSp} \rightarrow \text{SimpComp}$ dalla categoria degli spazi di tolleranza alla categoria dei complessi simpliciali finiti. La composizione nel funtore è compatibile con la composizione dei morfismi.

Dimostrazione. Si verifica direttamente che il morfismo identità $\text{id}_X: X \rightarrow X$ è un morfismo di spazi di tolleranza (che manda $x \sim_X x'$ in $x \sim_X x'$), che la composizione di 2 morfismi è ancora un morfismo (poiché $\varphi(x) \sim_Y \varphi(x')$ e $\psi(\varphi(x)) \sim_Z \psi(\varphi(x'))$ per morfismi $\varphi: X \rightarrow Y$ e $\psi: Y \rightarrow Z$), e che S manda identità in identità e composizioni in composizioni. $\square$

La struttura categoriale del complesso di Sorite ha implicazioni importanti per la teoria degli invarianti. In particolare, qualunque invariante che sia "funtoriale"—ovvero che si comporti in modo compatibile con i morfismi—è automaticamente un invariante degli spazi di tolleranza. I gruppi di omologia $H_k(S(X, \sim))$ sono invarianti funtoriali: un isomorfismo di spazi di tolleranza induce isomorfismi in omologia (ma non viceversa: l'omologia non classifica completamente i predicati in generale, come si discute nel Capitolo 11).

Proposizione 5 (Invarianza sotto isomorfismo). Se $\varphi: (X, \sim_X) \rightarrow (Y, \sim_Y)$ è un isomorfismo di spazi di tolleranza, allora $S(\varphi): S(X, \sim_X) \rightarrow S(Y, \sim_Y)$ è un isomorfismo di complessi simpliciali, e $\varphi^*: H_k(S(X, \sim_X)) \rightarrow H_k(S(Y, \sim_Y))$ è un isomorfismo di gruppi di omologia per ogni $k \geq 0$.

Dimostrazione. Se φ è biiettiva e preserva $\sim$ biunivocamente, allora $S(\varphi)$ è una biezione sui vertici che preserva l'essere-simplex biunivocamente (per la Proposizione 3 e la sua inversa $S(\varphi^{-1})$). Dunque $S(\varphi)$ è un isomorfismo simpliciale, e per la functorialità dell'omologia gli omomorfismi indotti sono isomorfismi. $\square$

3.5 Il complesso di Sorite come nuvola di dati topologica

Un'osservazione di carattere metodologico chiarisce il ruolo del complesso di Sorite nel contesto più ampio della TDA. La TDA (Topological Data Analysis) è un campo emerso negli anni 2000 che studia la "forma" dei dati usando strumenti topologici, con applicazioni in biologia, scienze dei materiali, analisi di reti sociali e molti altri domini.

Il paradigma standard della TDA è: dato un insieme di punti X in uno spazio metrico, costruire una filtrazione di complessi simpliciali dipendente da un parametro r (tipicamente la filtrazione di Vietoris–Rips), calcolare i gruppi di omologia di ogni complesso della filtrazione, e tracciare come queste strutture omologiche nascono e muoiono al variare di r . Il risultato è il diagramma di persistenza, che fornisce una firma topologica del dato che è stabile rispetto a perturbazioni.

Il complesso di Sorite è esattamente questa costruzione applicata al dominio del predicato vago: i "dati" sono le istanze del predicato (stature, lunghezze d'onda, ecc.), la metrica è la distanza di indistinguibilità percepita, e la filtrazione di Sorite è la filtrazione di Vietoris–Rips associata. La TDA fornisce dunque non solo strumenti matematici ma anche metodi computazionali consolidati per l'analisi dei predicati vaghi, come discusso nel Capitolo 13.

Capitolo 4. Il Paradosso di Sorite come Teorema

4.1 Assegnamenti di valori di verità e coerenza

L'analisi del paradosso di Sorite richiede di formalizzare con precisione le nozioni di "assegnamento di valori di verità" e di "coerenza" rispetto alla struttura di tolleranza. Queste nozioni, introdotte informalmente nella Sezione 1.1, ricevono ora una definizione rigorosa.

Definizione 8 (Assegnamento coerente). Dato uno spazio di tolleranza $(X, \sim)$ e un insieme V di valori di verità, un assegnamento coerente è una funzione $f: X \rightarrow V$ tale che $f(x) = f(y)$ ogni volta che $x \sim y$. La coerenza cattura il Principio di Tolleranza (P2): elementi indistinguibili ai fini del predicato devono ricevere lo stesso valore di verità.

Definizione 9 (Assegnamento monotono coerente). Se X è dotato di un ordinamento totale $\leq$, un assegnamento monotono coerente è una funzione $f: X \rightarrow \{\text{Vero}, \text{Falso}\}$ che soddisfa simultaneamente:

1. Coerenza: $f(x) = f(y)$ ogni volta che $x \sim y$.
2. Monotonicità non-decrescente: se $x \leq y$ e $f(x) = \text{Vero}$, allora $f(y) = \text{Vero}$.

La monotonicità formalizza l'intuizione che il predicato rispetti l'ordinamento: se n grani formano un mucchio, anche $n + 1$ grani lo formano; se una persona di altezza h è alta, anche una persona di altezza $h' > h$ lo è. Questa proprietà corrisponde alla premessa P2 del paradosso di Sorite interpretata in senso "ascendente": il predicato non può cessare di valere al crescere dell'argomento una volta che ha iniziato a valere.

Definizione 10 (Predicato con contrasto). Un assegnamento $f: X \rightarrow \{\text{Vero}, \text{Falso}\}$ ha contrasto se $f(\min(X)) = \text{Falso}$ e $f(\max(X)) = \text{Vero}$. Il contrasto cattura il fatto osservato che esistono esempi

paradigmatici di non-mucchio (1 grano) e di mucchio (10.000 grani), ossia che il predicato non è costantemente Vero né costantemente Falso sul dominio.

Il Teorema 1 affermerà che i 3 requisiti — coerenza, monotonicità, contrasto — sono congiuntamente insoddisfacibili quando il grafo di tolleranza è connesso. Prima di enunciarlo, è utile esaminare la struttura geometrica del problema.

4.2 Struttura geometrica del problema di assegnamento

Si consideri $X = \{x_1 < x_2 < \dots < x_n\}$ con la tolleranza $\sim$. Un assegnamento bivalente $f: X \rightarrow \{0, 1\}$ (dove 0 = Falso, 1 = Vero) è un punto nell'ipercubo $\{0, 1\}^n$. La condizione di coerenza $f(x_i) = f(x_j)$ per $x_i \sim x_j$ definisce un sistema di vincoli lineari (uguaglianze) che partiziona $\{0, 1\}^n$ in facce dell'ipercubo. La condizione di monotonicità $f(x_i) \leq f(x_j)$ per $i \leq j$ è un ulteriore sistema di vincoli lineari (disuguaglianze). La condizione di contrasto $f(x_1) = 0$ e $f(x_n) = 1$ fissa 2 coordinate.

Il Teorema 1 dice che l'intersezione di questi 3 sistemi di vincoli con $\{0, 1\}^n$ è vuota quando $G(X, \sim)$ è connesso. La dimostrazione che segue è elementare, ma il risultato ha una profondità che si manifesta nelle sue generalizzazioni ai valori graduati (Teorema 2) e alle logiche a gradi (Capitolo 5).

4.3 Il teorema principale di impossibilità

Teorema 1 (Il paradosso di Sorite come teorema). Sia $(X, \leq)$ un insieme totalmente ordinato con minimo $\min(X)$ e massimo $\max(X)$, e sia $\sim$ una relazione di tolleranza su X tale che il grafo di tolleranza $G(X, \sim)$ sia connesso. Allora non esiste nessun assegnamento $f: X \rightarrow \{\text{Vero}, \text{Falso}\}$ che sia simultaneamente coerente, monotono e con contrasto.

Dimostrazione. Si assume per assurdo che esista un tale f . Poiché f è coerente, si ha $f(x) = f(y)$ ogni volta che $x \sim y$; poiché $G(X, \sim)$ è connesso, esiste per ogni coppia $x, y \in X$ un cammino $x = z_0, z_1, \dots, z_m = y$ con $z_i \sim z_{i+1}$ per ogni i . Per transitività della relazione di uguaglianza dei valori di f lungo il cammino: $f(z_0) = f(z_1) = \dots = f(z_m)$, dunque $f(x) = f(y)$. Poiché questo vale per ogni coppia $x, y \in X$ (il grafo è connesso), f è costantemente uguale su X . In particolare $f(\min(X)) = f(\max(X))$, che contraddice il contrasto $f(\min(X)) = \text{Falso} \neq \text{Vero} = f(\max(X))$. $\square$

La dimostrazione è elementare, ma è utile evidenziare dove entra ciascuna delle 3 ipotesi del teorema. La connessione di $G(X, \sim)$ garantisce che f debba essere costante sull'intero X se è coerente. Il contrasto richiede che f assuma 2 valori diversi su X . La monotonicità, sorprendentemente, non entra nella dimostrazione del Teorema 1 come enunciato: la contraddizione si ottiene già dalla coerenza e dal contrasto. La monotonicità diventa rilevante

quando si generalizza agli assegnamenti graduati, perché senza di essa un assegnamento "oscillante" potrebbe in linea di principio avere $f(\min(X)) = 0$, $f(\max(X)) = 1$ e tuttavia essere ε -coerente per ε sufficientemente grande.

Il Teorema 1 mostra che i 3 requisiti seguenti sono congiuntamente insoddisfacibili quando $G(X, \sim)$ è connesso:

- (R1) Principio di tolleranza: elementi tolleranti ricevono lo stesso valore di verità.
- (R2) Monotonicità: il valore di verità non decresce nell'ordinamento.
- (R3) Contrasto: il minimo ha valore Falso e il massimo ha valore Vero.

Queste 3 condizioni corrispondono esattamente alle premesse del paradosso di Sorite nella sua formulazione filosofica: R1 è P2, R3 è P1 (con l'aggiunta del fatto che esistono mucchi), e la monotonicità è implicita nell'uso di P2 come regola d'induzione. Il Teorema 1 dice che qualunque risposta filosofica al paradosso deve negare almeno 1 dei 3 requisiti: l'epistemicismo nega R1 (almeno per la "vera" relazione di tolleranza), la supervalutazione nega R3 (dichiarando che l'enunciato "max(X) è mucchio" non è supervero per una certa precisazione), e le logiche fuzzy abbandonano il bivalentismo che rende sensati R1, R2, R3 come formulati.

Corollario 1.1. Un assegnamento coerente $f: X \rightarrow \{\text{Vero}, \text{Falso}\}$ con contrasto ($f(\min(X)) = \text{Falso}$, $f(\max(X)) = \text{Vero}$) esiste se e solo se $\min(X)$ e $\max(X)$ appartengono a componenti connesse diverse di $G(X, \sim)$.

Dimostrazione. ($\Rightarrow$) Se esistono tali f coerente con contrasto, allora $f(\min(X)) \neq f(\max(X))$, dunque $\min(X)$ e $\max(X)$ non possono essere nella stessa componente connessa (perché in una componente connessa tutti i vertici hanno lo stesso valore di f , per coerenza e connessione). ($\Leftarrow$) Se $\min(X)$ e $\max(X)$ sono in componenti diverse $C_{\min}$ e $C_{\max}$, si definisce $f(x) = \text{Falso}$ per $x \in C_{\min}$ e $f(x) = \text{Vero}$ per $x \in C_{\max}$. Per tutti le altre componenti (se esistono), si assegna Falso o Vero in modo da rispettare la monotonicità: le componenti "sotto" $C_{\min}$ nell'ordinamento ricevono Falso, quelle "sopra" $C_{\max}$ ricevono Vero, e quelle intermedie ricevono il valore della componente immediatamente precedente. La coerenza è soddisfatta per costruzione (f è costante su ogni componente). Il contrasto è soddisfatto per la scelta di f su $C_{\min}$ e $C_{\max}$. La monotonicità richiede che le componenti siano "separate" nell'ordinamento, il che è garantito dal fatto che $G(X, \sim)$ è disconnesso (la disconnessione separa X in blocchi ordinati). $\square$

Corollario 1.2 (Localizzazione della vaghezza). La vaghezza del predicato — nel senso dell'impossibilità di un assegnamento coerente con contrasto — è completamente determinata dalla struttura di connessione del grafo $G(X, \sim)$. Precisamente, il predicato ammette un assegnamento coerente con contrasto se e solo se esiste un taglio del grafo $G(X, \sim)$ che separa $\min(X)$ da $\max(X)$: una partizione $X = S \cup T$ con $\min(X) \in S$, $\max(X) \in T$ e nessun arco di $G(X, \sim)$ tra S e T .

Dimostrazione. Un taglio che separa $\min(X)$ da $\max(X)$ è equivalente a $\min(X)$ e $\max(X)$ che appartengono a componenti connesse diverse del grafo ottenuto rimuovendo certi archi. Ma i tagli rilevanti sono quelli che rimuovono tutti gli archi tra le componenti: questo è esattamente la disconnessione tra le componenti contenenti $\min(X)$ e $\max(X)$. Il corollario segue dal Corollario 1.1. $\square$

Il Corollario 1.2 suggerisce una misura qualitativa della vaghezza: il predicato è tanto più "vago" quanto più difficile è tagliare il grafo per separare $\min(X)$ da $\max(X)$. Questa difficoltà è catturata dal numero di archi che occorre rimuovere per disconnettere $\min(X)$ da $\max(X)$ — la connettività tra $\min(X)$ e $\max(X)$ nel senso di Menger — e dalla loro distribuzione nel grafo. Per misure quantitative più precise, il Capitolo 5 introduce il numero di risoluzione ϵ^* .

4.4 Estensione alle logiche a gradi

Il Teorema 1 è formulato per valori di verità bivalenti $V = \{\text{Vero}, \text{Falso}\}$. Nella tradizione delle logiche fuzzy, i valori di verità sono nell'intervallo $[0, 1]$. La domanda naturale è: cosa succede al paradosso di Sorite quando si ammettono valori intermedi?

Intuitivamente, si potrebbe pensare che consentire valori in $[0, 1]$ risolva il problema: si potrebbe assegnare 0 al minimo, 1 al massimo, e valori intermedi agli elementi nella zona di confine, con una transizione graduale. Ma il problema è che la transizione deve essere "coerente" con la struttura di tolleranza: elementi indistinguibili devono ricevere valori simili. Quanto simili? Questo è precisamente il problema del numero di risoluzione ϵ^* .

Definizione 11 (Assegnamento ϵ -coerente). Dato uno spazio di tolleranza $(X, \sim)$ e un parametro $\epsilon \geq 0$, un assegnamento ϵ -coerente è una funzione $f: X \rightarrow [0, 1]$ tale che $|f(x) - f(y)| \leq \epsilon$ ogni volta che $x \sim y$.

Per $\epsilon = 0$ si recupera la coerenza esatta della Definizione 8 (con $V = [0, 1]$): un assegnamento 0-coerente assegna lo stesso valore a tutti gli elementi indistinguibili. Per $\epsilon = 1$, ogni funzione $f: X \rightarrow [0, 1]$ è banalmente 1-coerente, poiché $|f(x) - f(y)| \leq 1$ per ogni $x, y \in [0, 1]$ per definizione.

Teorema 2 (Impossibilità graduata). Sia $(X, \sim, \leq)$ uno spazio di tolleranza totalmente ordinato con $G(X, \sim)$ connesso, e sia $L_{\text{mono_min}}$ la lunghezza del cammino monotono minimo da $\min(X)$ a $\max(X)$ in $G(X, \sim)$ (definito nel Capitolo 5). Per ogni $\epsilon < 1/L_{\text{mono_min}}$, non esiste nessun assegnamento ϵ -coerente monotono $f: X \rightarrow [0, 1]$ con $f(\min(X)) = 0$ e $f(\max(X)) = 1$.

Dimostrazione. Segue dal Teorema 3 del Capitolo 5, che afferma $\epsilon^*(\sim) = 1/L_{\text{mono_min}}$. Per la definizione di ϵ^* come infimo, ogni assegnamento ϵ -coerente monotono con condizioni agli estremi richiede $\epsilon \geq \epsilon^*$. $\square$

Il Teorema 2 è la versione graduata del paradosso di Sorite: anche consentendo valori di verità continui in $[0, 1]$, il costo minimo di coerenza ε^* è strettamente positivo quando $G(X, \sim)$ è connesso. Non si può ottenere una transizione "infinitamente graduata" ($\varepsilon \rightarrow 0$) da Falso a Vero: c'è un costo minimo positivo di incoerenza, che dipende dalla lunghezza del cammino monotono più corto nel grafo di tolleranza.

4.5 La struttura dei 3 corni del dilemma

Alla luce del Teorema 1 e del Teorema 2, è possibile descrivere con precisione matematica la struttura del "dilemma" filosofico del paradosso di Sorite.

Dato uno spazio di tolleranza $(X, \sim, \leq)$ con $G(X, \sim)$ connesso, le 3 opzioni per il filosofo sono:

Opzione A (Epistemica): si abbandona R1 sostenendo che la relazione di tolleranza "vera" $\sim_{\text{vero}}$ è diversa da quella osservata $\sim_{\text{obs}}$, e che $G(X, \sim_{\text{vero}})$ è disconnesso. Il costo è la distanza bottleneck $d_B(B(\sim_{\text{obs}}), B(\sim_{\text{vero}}))$ tra i sorites barcode delle 2 relazioni di tolleranza, che misura quanto devono "discordare" le 2 relazioni.

Opzione B (Supervaluazionista): si abbandona R3 sostenendo che l'enunciato " $\max(X)$ soddisfa il predicato" non è supervero. Il costo è la distanza tra la struttura di tolleranza e il predicato nitido più vicino nel senso di Hausdorff: si misura quanto archi del grafo occorra aggiungere o rimuovere per ottenere la disconnessione tra $\min(X)$ e $\max(X)$.

Opzione C (Graduata): si abbandona il bivalentismo e si lavora con $V = [0, 1]$. Il costo è il numero di risoluzione $\varepsilon^* = 1/L_{\text{mono_min}}$, che misura il grado minimo di "incoerenza graduata" da accettare per ottenere un assegnamento con contrasto.

La matematica non dice quale opzione scegliere; mostra soltanto che le 3 opzioni hanno costi misurabili e che il costo è zero (nel senso appropriato per ciascuna opzione) se e solo se $G(X, \sim)$ è disconnesso con $\min(X)$ e $\max(X)$ in componenti diverse. In quel caso, il predicato non è vago (è "nitido") e il paradosso di Sorite non si pone.

Capitolo 5. Il Numero di Risoluzione

5.1 Motivazione e definizione

Il Teorema 1 dimostra che la coerenza esatta è incompatibile con il contrasto quando $G(X, \sim)$ è connesso. Il Teorema 2 generalizza questo risultato alle logiche a gradi, mostrando che anche la ε -coerenza per ε piccolo è incompatibile con il contrasto. La domanda quantitativa naturale è: qual è il valore esatto di ε^* al di sotto del quale l'incompatibilità vale, e al di sopra del quale l'assegnamento con contrasto è possibile?

Definizione 12 (Numero di risoluzione). Per uno spazio di tolleranza $(X, \sim, \leq)$ con minimo $\min(X)$ e massimo $\max(X)$, il numero di risoluzione è:

$$\varepsilon^*(\sim) = \inf\{\varepsilon \geq 0 : \text{esiste un assegnamento } \varepsilon\text{-coerente monotono } f: X \rightarrow [0, 1] \text{ con } f(\min(X)) = 0 \text{ e } f(\max(X)) = 1\}$$

L'infimo è su tutti i valori non negativi di ε per i quali esiste un assegnamento del tipo richiesto. Il numero ε^* misura il costo minimo di risoluzione del paradosso: la quantità minima di "incoerenza graduata" che bisogna accettare perché esista un assegnamento accettabile che distingue gli estremi.

Il nome "numero di risoluzione" riflette l'idea che ε^* quantifichi la "risoluzione minima" con cui il predicato può essere valutato: se si accetta un'incertezza di ε unità nel confronto tra valori di elementi indistinguibili, il paradosso è "risolto" (l'assegnamento con contrasto esiste) per $\varepsilon \geq \varepsilon^*$, e "irrisolto" per $\varepsilon < \varepsilon^*$.

Proposizione 6 (Proprietà elementari di ε^*). Valgono i seguenti fatti:

- (a) $\varepsilon^*(\sim) \geq 0$ sempre, e $\varepsilon^*(\sim) = 0$ se e solo se $\min(X)$ e $\max(X)$ appartengono a componenti connesse diverse di $G(X, \sim)$.
- (b) $\varepsilon^*(\sim) \leq 1$ sempre.
- (c) $\varepsilon^*(\sim) > 0$ se e solo se $G(X, \sim)$ è connesso.

Dimostrazione di (a). L'infimo è su una quantità non negativa, dunque $\varepsilon^* \geq 0$. Se $\min(X)$ e $\max(X)$ sono in componenti diverse, la funzione che assegna 0 alla componente di $\min(X)$ e 1 alla componente di $\max(X)$ è 0-coerente (costante su ogni componente), dunque $\varepsilon^* = 0$. Viceversa, se sono nella stessa componente, ogni assegnamento 0-coerente è costante sull'intera componente (per la Proposizione 2), dunque $f(\min(X)) = f(\max(X))$, impossibile con il contrasto; pertanto $\varepsilon^* > 0$.

Dimostrazione di (b). La funzione lineare $f(x_i) = (i - 1)/(n - 1)$ per $x_i \in X = \{x_1 < \dots < x_n\}$ è monotona per costruzione, soddisfa $f(x_1) = 0$ e $f(x_n) = 1$. Per qualunque coppia tollerante $x_i \sim x_j$ con $i \leq j$, $|f(x_i) - f(x_j)| = (j - i)/(n - 1) \leq (n - 1)/(n - 1) = 1$. Dunque f è 1-coerente, e $\varepsilon^*(\sim) \leq 1$.

Dimostrazione di (c). Dalla (a). $\square$

La Proposizione 6 caratterizza il numero di risoluzione nel caso estremi. Il caso interessante è il caso intermedio: $G(X, \sim)$ connesso, $\varepsilon^*(\sim) \in (0, 1]$. La teoria che segue calcola ε^* esplicitamente in funzione della struttura del grafo.

5.2 Cammini monotoni nel grafo di tolleranza

La formula esatta per $\varepsilon^*(\sim)$ richiede la nozione di cammino monotono, che formalizza il percorso "ordinato" nel grafo di tolleranza da un estremo all'altro.

Definizione 13 (Cammino monotono). In uno spazio di tolleranza $(X, \sim, \leq)$ con $X = \{x_1 < x_2 < \dots < x_n\}$ totalmente ordinato, un cammino monotono da $\min(X) = x_1$ a $\max(X) = x_n$ è una sequenza $y_0, y_1, \dots, y_L$ di elementi di X con:

1. $y_0 = x_1 = \min(X)$ e $y_L = x_n = \max(X)$.
2. $y_i \sim y_{i+1}$ per ogni $i = 0, 1, \dots, L - 1$ (ogni passo percorre un arco del grafo di tolleranza).
3. $y_i \leq y_{i+1}$ per ogni $i = 0, 1, \dots, L - 1$ (il cammino è strettamente non-decrescente nell'ordinamento).

La lunghezza del cammino è L (il numero di archi percorsi). Il cammino non è richiesto di passare per tutti i vertici di X , ma soltanto di muoversi "in avanti" nell'ordinamento ad ogni passo. Si denota con $L_{\text{mono_min}}$ la lunghezza del cammino monotono minimo da x_1 a x_n .

Lemma 1 (Esistenza del cammino monotono). Se $G(X, \sim)$ è connesso e $(X, \leq)$ è totalmente ordinato, allora esiste almeno un cammino monotono da $\min(X)$ a $\max(X)$, e quindi $L_{\text{mono_min}}$ è ben definito.

Dimostrazione. Poiché $G(X, \sim)$ è connesso, esiste un cammino (non necessariamente monotono) $z_0 = x_1, z_1, \dots, z_m = x_n$ in $G(X, \sim)$. Si applica un processo di "raddrizzamento monotono". Si considera la lunghezza del segmento non-monotono più lungo del cammino: se il cammino è già monotono, si è finito. Altrimenti, esiste un indice i tale che $z_{i+1} < z_i$ (il cammino "retrocede"). Si possono distinguere due casi.

Caso 1: z_{i-1} e z_{i+1} sono adiacenti nel grafo ($z_{i-1} \sim z_{i+1}$). Allora si può "cortocircuitare" il passo non-monotono sostituendo il sottocammino $\dots, z_{i-1}, z_i, z_{i+1}, \dots$ con $\dots, z_{i-1}, z_{i+1}, \dots$. Il cammino risultante ha 1 passo in meno e 1 passo non-monotono in meno.

Caso 2: z_{i-1} e z_{i+1} non sono adiacenti nel grafo. In questo caso non si può cortocircuitare direttamente, ma si può ricorrere all'induzione: il sottocammino da z_{i-1} a z_{i+1} (che passa per z_i) ha lunghezza 2, e $z_{i-1} \leq z_i$ ma $z_{i+1} < z_i$. Se $z_{i-1} \leq z_{i+1}$, il passo da z_{i-1} a z_{i+1} esiste ma non è un arco del grafo. Si deve trovare un cammino diverso nel grafo da z_{i-1} a z_{i+1} che sia monotono.

Il punto sottile è che questo processo non sempre termina immediatamente, ma si può dimostrare la terminazione per induzione sull'indice di "inversione": il numero di coppie (i, j) con $i < j$ ma $z_i > z_j$ (inversioni nell'ordinamento). Ogni passo del raddrizzamento riduce strettamente il numero di inversioni se riduce la lunghezza del cammino. Poiché il numero di inversioni è non negativo e intero, il processo termina dopo un numero finito di passi.

Per mostrare che il processo termina con un cammino monotono (non semplicemente con un cammino più corto), si usa la connessione di $G(X, \sim)$: qualunque passo retrogrado può essere "aggirato" perché il grafo è connesso, e dunque esiste sempre un cammino monotono alternativo. L'argomento formale richiede di mostrare che il DAG $G^\rightarrow(X, \sim)$ (descritto nel Lemma 2) ha un cammino da x_1 a x_n , il che segue dalla connessione di $G(X, \sim)$ tramite l'argomento di raddrizzamento. $\square$

Lemma 2 (Il DAG monotono). Si orienta il grafo $G(X, \sim)$ mettendo un arco diretto da x_i a x_j se e solo se $x_i \sim x_j$ e $x_i < x_j$ (nell'ordinamento totale su X). Il grafo orientato risultante, denotato $G^\rightarrow(X, \sim)$, è un DAG (grafo orientato aciclico) perché l'ordinamento totale garantisce che non esistano cicli diretti: se ci fosse un ciclo diretto $x_{i_1} \rightarrow x_{i_2} \rightarrow \dots \rightarrow x_{i_k} \rightarrow x_{i_1}$, l'ordinamento implicherebbe $x_{i_1} < x_{i_2} < \dots < x_{i_k} < x_{i_1}$, contraddizione.

La lunghezza $L_{\text{mono_min}}$ è quindi la lunghezza del cammino diretto minimo da x_1 a x_n in $G^\rightarrow(X, \sim)$, che si calcola efficientemente con la programmazione dinamica su DAG.

Verifica. Ogni cammino monotono in $G(X, \sim)$ corrisponde biunivocamente a un cammino diretto in $G^\rightarrow(X, \sim)$ (poiché la monotonia $y_i \leq y_{i+1}$ con $y_i \sim y_{i+1}$ equivale all'arco diretto $y_i \rightarrow y_{i+1}$ in $G^\rightarrow$). L'identificazione è completa. $\square$

5.3 Il teorema principale: la formula per ε^*

Teorema 3 (Formula per il numero di risoluzione). Per uno spazio di tolleranza $(X, \sim, \leq)$ con $G(X, \sim)$ connesso, vale:

$$\varepsilon^*(\sim) = 1 / L_{\text{mono_min}}$$

dove $L_{\text{mono_min}}$ è la lunghezza del cammino monotono minimo da $\min(X)$ a $\max(X)$ in $G^+(X, \sim)$.

Dimostrazione. La dimostrazione si divide in 2 parti: si mostra prima che $\varepsilon^*(\sim) \leq 1/L_{\text{mono_min}}$ costruendo esplicitamente un assegnamento $(1/L_{\text{mono_min}})$ -coerente, poi che $\varepsilon^*(\sim) \geq 1/L_{\text{mono_min}}$ mostrando che ogni assegnamento ε -coerente con contrasto deve avere $\varepsilon \geq 1/L_{\text{mono_min}}$.

Parte 1: $\varepsilon(\sim) \leq 1/L_{\text{mono_min}}.$ *

Sia $y_0 = x_1, y_1, \dots, y_L = x_n$ il cammino monotono minimo di lunghezza $L = L_{\text{mono_min}}$. Si definisce la funzione gradino $f: X \rightarrow [0, 1]$ ponendo, per ogni $x \in X$:

$$f(x) = j(x) / L, \text{ dove } j(x) = \max \{j : y_j \leq x\}.$$

In parole, f assegna a x il valore j/L dove j è l'indice del nodo del cammino minimo più grande che non supera x . Si verifica che f soddisfa le tre proprietà richieste.

— Monotonicità: se $x \leq z$ allora ogni $y_j \leq x$ soddisfa $y_j \leq z$, dunque $j(x) \leq j(z)$ e $f(x) \leq f(z)$. ✓

— Condizioni agli estremi: $f(x_1) = j(x_1)/L = 0/L = 0$ (poiché $y_0 = x_1$ e nessun y_j con $j > 0$ soddisfa $y_j \leq x_1$) e $f(x_n) = j(x_n)/L = L/L = 1$ (poiché $y_L = x_n$). ✓

— $(1/L)$ -coerenza: per ogni coppia tollerante $x_i \sim x_j$ con $i \leq j$ si ha $j(x_i) \leq j(x_j)$ (per monotonicità), dunque $|f(x_i) - f(x_j)| = (j(x_j) - j(x_i))/L$. Si mostra che $j(x_j) - j(x_i) \leq L$.

Si supponga per assurdo che $j(x_j) - j(x_i) \geq 2$. Posto $a = j(x_i)$ e $b = j(x_j)$, si costruisce il seguente cammino monotono da x_1 a x_n : si percorre il cammino minimo originale da x_1 fino a y_a , poi si percorre l'arco diretto $x_i \rightarrow x_j$ (arco che esiste per ipotesi $x_i \sim x_j$ con $x_i < x_j$), infine si riprende il cammino minimo originale da y_{b+1} fino a x_n . Poiché $y_a \leq x_i < x_j \leq y_{b+1}$ e gli archi $y_a \rightarrow x_i$ e $x_j \rightarrow y_{b+1}$ sono sostituiti da un unico arco $x_i \rightarrow x_j$, la lunghezza del cammino alternativo è:

$$a + 1 + (L - b - 1) = L - (b - a - 1) \leq L - 1 < L,$$

poiché $b - a \geq 2$ implica $b - a - 1 \geq 1$. Questo contraddice la minimalità di L . Dunque $j(x_j) - j(x_i) \leq 1$, e pertanto $|f(x_i) - f(x_j)| \leq 1/L$. ✓

Le tre proprietà sono dunque soddisfatte, il che completa la dimostrazione che $\varepsilon^*(\sim) \leq 1/L_{\text{mono_min}}$. □

Parte 2: $\varepsilon(\sim) \geq 1/L_{\text{mono_min}}$.*

Sia $f: X \rightarrow [0, 1]$ un qualunque assegnamento ε -coerente monotono con $f(x_1) = 0$ e $f(x_n) = 1$. Si vuole dimostrare che $\varepsilon \geq 1/L_{\text{mono_min}}$.

Sia $y_0 = x_1, y_1, \dots, y_L = x_n$ un qualunque cammino monotono da x_1 a x_n di lunghezza L . Poiché f è monotona e il cammino è monotono ($y_i \leq y_{i+1}$ per ogni i), si ha $f(y_0) \leq f(y_1) \leq \dots \leq f(y_L)$. La somma telescopica fornisce:

$$1 = f(y_L) - f(y_0) = \sum_{i=0}^{L-1} [f(y_{i+1}) - f(y_i)]$$

Ogni termine $f(y_{i+1}) - f(y_i)$ è non-negativo (per monotonia) e soddisfa $f(y_{i+1}) - f(y_i) \leq \varepsilon$ (per ε -coerenza, poiché $y_i \sim y_{i+1}$). Dunque:

$$1 = \sum_{i=0}^{L-1} [f(y_{i+1}) - f(y_i)] \leq L \cdot \varepsilon$$

il che dà $\varepsilon \geq 1/L$. Poiché questo vale per qualunque cammino monotono di lunghezza L , e in particolare per il cammino monotono minimo di lunghezza $L_{\text{mono_min}}$, si ha:

$$\varepsilon \geq 1/L_{\text{mono_min}}$$

Poiché questo vale per ogni assegnamento ε -coerente monotono con contrasto, si ha $\varepsilon^*(\sim) \geq 1/L_{\text{mono_min}}$.

Conclusione. Dalle 2 parti segue $\varepsilon^*(\sim) = 1/L_{\text{mono_min}}$. □

5.4 Analisi della Parte 2: perché il cammino monotono minimo è quello rilevante

La Parte 2 della dimostrazione usa l'argomento telescopico su un qualunque cammino monotono per dare un lower bound su ε . Il lower bound ottenuto usando un cammino di lunghezza L è $1/L$.

Dunque il lower bound più forte si ottiene usando il cammino più lungo, non il più corto. Ma la formula dice $\varepsilon^* = 1/L_{\text{mono_min}}$, ossia si usa il cammino più corto.

La spiegazione è la seguente: si cerca il valore minimo di ε per cui esiste un assegnamento ammissibile. La Parte 2 dice che per ogni assegnamento ammissibile e per ogni cammino monotono di lunghezza L , si ha $\varepsilon \geq 1/L$. Il lower bound più forte su ε^* (l'infimo su tutti gli assegnamenti) si ottiene scegliendo il lower bound più forte, che è $\varepsilon \geq 1/L_{\text{mono_min}}$ (poiché $L_{\text{mono_min}}$ è il cammino più corto, $1/L_{\text{mono_min}}$ è il massimo tra $1/L$ al variare di L sui cammini monotoni).

In sintesi: la Parte 1 costruisce un assegnamento $(1/L_{\text{mono_min}})$ -coerente, la Parte 2 mostra che nessun assegnamento può avere $\varepsilon < 1/L_{\text{mono_min}}$. Insieme danno l'uguaglianza esatta.

5.5 Il rapporto con la bandwidth costretta e le discrepanze

Proposizione 7 (Disuguaglianza fondamentale). Per ogni spazio di tolleranza $(X, \sim, \leq)$ con $G(X, \sim)$ connesso e $n = |X|$, vale:

$$1/(n-1) \leq \varepsilon^*(\sim) = 1/L_{\text{mono_min}} \leq bw_{\leq}(\sim)/(n-1)$$

Dimostrazione del bound inferiore: ogni cammino monotono da x_1 a x_n deve "coprire" almeno $n-1$ posizioni ordinali, e ogni passo copre al massimo $bw_{\leq}(\sim)$ posizioni. Dunque $L_{\text{mono_min}} \geq (n-1)/bw_{\leq}(\sim)$, e $\varepsilon^*(\sim) = 1/L_{\text{mono_min}} \leq bw_{\leq}(\sim)/(n-1)$. Per il bound inferiore, ogni cammino monotono ha almeno 1 passo, quindi $L_{\text{mono_min}} \geq 1$, ma in realtà il cammino più corto non può avere meno di $\lceil (n-1)/bw_{\leq} \rceil$ passi (poiché deve coprire $n-1$ posizioni ordinali con passi di lunghezza massima $bw_{\leq}$), dunque $L_{\text{mono_min}} \geq \lceil (n-1)/bw_{\leq} \rceil \geq (n-1)/bw_{\leq}$. Tuttavia il bound inferiore $1/(n-1)$ si ottiene banalmente dal fatto che $L_{\text{mono_min}} \leq n-1$ (il cammino più lungo su n vertici distinti ha $n-1$ passi). $\square$

Proposizione 8 (Coincidenza per tolleranze uniformi). Per la tolleranza di larghezza uniforme k su $X = \{1, \dots, n\}$, si ha $\varepsilon^*(\sim) = k/(n-1)$.

Dimostrazione. Il cammino monotono minimo da 1 a n nella tolleranza uniforme k è la sequenza $1, 1+k, 1+2k, \dots, 1+\lceil (n-1)/k \rceil \cdot k$. Se k divide $n-1$, la lunghezza è esattamente $(n-1)/k$ e $\varepsilon^*(\sim) = k/(n-1)$. In generale, $L_{\text{mono_min}} = \lceil (n-1)/k \rceil$ e $\varepsilon^*(\sim) = 1/\lceil (n-1)/k \rceil$, che per k che divide $n-1$ dà esattamente $k/(n-1)$. Inoltre $bw_{\leq}(\sim) = k$ per definizione, dunque $\varepsilon^*(\sim) = bw_{\leq}(\sim)/(n-1)$ in questo caso: la disuguaglianza della Proposizione 7 diventa un'uguaglianza. $\square$

Proposizione 9 (Discrepanza massima tra ε e $bw_{\leq}/(n-1)$). * Esistono spazi di tolleranza per cui il rapporto $\varepsilon^*(\sim) / (bw_{\leq}(\sim)/(n-1))$ può essere arbitrariamente piccolo (precisamente, dell'ordine di $1/(n-1)$).

Dimostrazione (costruzione esplicita). Sia $X = \{1, 2, \dots, n\}$ con la relazione: $i \sim j$ sse $|i - j| \leq 1$ (adiacenza nella catena) oppure $\{i, j\} = \{1, n\}$ (arco "lungo" tra gli estremi). Il grafo $G(X, \sim)$ è la catena $1 - 2 - \dots - n$ con 1 arco aggiuntivo $1 \sim n$. La bandwidth costretta è $bw_{\leq} = n - 1$ (la coppia $\{1, n\}$ ha distanza ordinale $n - 1$). Il cammino monotono minimo è però la catena $1, 2, 3, \dots, n$ di lunghezza $n - 1$ (l'arco $1 \sim n$ va "all'indietro" da n verso 1 in senso ordinale, dunque non è usabile in un cammino monotono). Quindi $L_{\text{mono_min}} = n - 1$ e $\varepsilon^*(\sim) = 1/(n - 1)$. Il rapporto è $\varepsilon^*(\sim) / (bw_{\leq}/(n - 1)) = (1/(n - 1)) / ((n - 1)/(n - 1)) = 1/(n - 1)^2 \rightarrow 0$ per $n \rightarrow \infty$. $\square$

La Proposizione 9 rivela un fatto importante: la presenza di archi "lunghi" nel grafo di tolleranza (alta bandwidth) non garantisce un numero di risoluzione basso. Ciò che conta per ε^* è se gli archi lunghi si trovano nella direzione "giusta" — nella direzione monotona dall'estremo inferiore a quello superiore — oppure "controcorrente". Archi che connettono l'estremo superiore all'estremo inferiore (come l'arco $1 \sim n$ nell'esempio) hanno alta bandwidth ma zero utilità per il cammino monotono.

Questa osservazione ha una conseguenza semantica: due predicati con la stessa bandwidth costretta possono avere numeri di risoluzione molto diversi, a seconda di come sono distribuite le relazioni di tolleranza rispetto all'ordinamento semantico. Un predicato con alcune relazioni di tolleranza "a lungo raggio" in direzione crescente (es. "grande" vs "piccolo" per oggetti di dimensioni molto diverse ma confrontabili) ha un ε^* basso; un predicato con relazioni di tolleranza a lungo raggio in direzione decrescente (es. una forma di "sommiglianza regressiva") ha un ε^* alto nonostante la stessa bandwidth.

5.6 Unicità e struttura degli assegnamenti ottimali

Proposizione 10 (Unicità sui nodi del cammino minimo). Se il cammino monotono minimo $y_0 = x_1, y_1, \dots, y_L = x_n$ da $\min(X)$ a $\max(X)$ è unico, allora ogni assegnamento f che realizza il minimo $\varepsilon^* = 1/L$ (ovvero è $(1/L)$ -coerente, monotono, e soddisfa $f(x_1) = 0$, $f(x_n) = 1$) soddisfa $f(y_j) = j/L$ per ogni $j = 0, 1, \dots, L$.

Dimostrazione. Per la Parte 2 della dimostrazione del Teorema 3, applicata al cammino minimo $y_0, \dots, y_L$:

$$1 = \sum_{j=0}^{L-1} [f(y_{j+1}) - f(y_j)]$$

Ogni termine è non-negativo (monotonia) e $\leq \varepsilon^* = 1/L$ (ε^* -coerenza). Poiché la somma di L termini, ciascuno $\leq 1/L$, è uguale a $1 = L \cdot (1/L)$, ogni termine deve essere esattamente $1/L$. Dunque $f(y_{j+1}) - f(y_j) = 1/L$ per ogni j , il che dà $f(y_j) = j/L$ per ogni j . $\square$

Proposizione 11 (Struttura degli assegnamenti per gli elementi fuori dal cammino). Per gli elementi $x \in X$ che non appartengono al cammino monotono minimo, il valore $f(x)$ può variare in un intervallo chiuso $[a(x), b(x)] \subseteq [0, 1]$ determinato dalle relazioni di tolleranza e di ordinamento di x con gli altri elementi.

Dimostrazione. Ogni elemento x non nel cammino minimo deve soddisfare:

- Monotonicità: $f(x) \leq f(x')$ per tutti gli $x' \geq x$ nel dominio.
- ϵ^* -coerenza: $|f(x) - f(z)| \leq \epsilon^*$ per tutti gli $z \sim x$.

Questi vincoli definiscono l'intervallo:

$$a(x) = \max\{f(z) - \epsilon^* : z \sim x\} \vee \max\{f(x') : x' \leq x, x' \in X\} \quad b(x) = \min\{f(z) + \epsilon^* : z \sim x\} \wedge \min\{f(x') : x' \geq x, x' \in X\}$$

dove max e min sono presi sulle relazioni di tolleranza e di ordine pertinenti. Poiché la costruzione della Parte 1 fornisce un assegnamento ammissibile (che assegna un valore a x), l'intervallo $[a(x), b(x)]$ è non vuoto. $\square$

La Proposizione 11 mostra che gli assegnamenti ottimali non sono in generale unici: c'è un grado di libertà per ogni elemento fuori dal cammino minimo. Questo grado di libertà ha un'interpretazione semantica: gli elementi fuori dal "percorso critico" del paradosso hanno valori di verità che non sono determinati dalla struttura di tolleranza, ma oscillano in un intervallo compatibile con i vincoli. Sono gli elementi meno "critici" ai fini del paradosso.

5.7 Calcolo di ϵ^* e complessità computazionale

L'algoritmo per calcolare ϵ^* segue direttamente dalla struttura del DAG $G^{\rightarrow}(X, \sim)$:

Algoritmo per $\epsilon(\sim)^*$:

1. Ordinare X : $x_1 < x_2 < \dots < x_n$.
2. Inizializzare $d(x_1) = 0$ (lunghezza del cammino minimo da x_1 a x_1 , che è 0).
3. Per $i = 2, \dots, n$: calcolare $d(x_i) = 1 + \min\{d(x_j) : x_j \sim x_i, j < i\}$. Se non esiste $j < i$ con $x_j \sim x_i$, porre $d(x_i) = +\infty$.

4. $L_mono_min = d(x_n)$ e $\varepsilon^*(\sim) = 1/d(x_n)$.

Proposizione 12 (Complessità del calcolo). Il calcolo di $\varepsilon^*(\sim)$ richiede tempo $O(|X| + |E|)$ dove $|E|$ è il numero di coppie tolleranti.

Dimostrazione. Il passo 3 dell'algoritmo è esattamente la ricerca del cammino minimo in un DAG con tecnica di programmazione dinamica. Poiché il DAG $G^{\rightarrow}(X, \sim)$ è già in ordine topologico (l'ordinamento di X è un ordinamento topologico del DAG, poiché tutti gli archi vanno da indice minore a indice maggiore), il calcolo si esegue in ordine crescente di i , e ogni arco (x_j, x_i) è visitato esattamente 1 volta. La complessità totale è $O(|V| + |E|) = O(|X| + |E|)$. $\square$

Proposizione 13 (NP-difficoltà dell'ottimizzazione sull'ordinamento). Il problema "dato uno spazio di tolleranza $(X, \sim)$ senza ordinamento prefissato, trovare l'ordinamento totale $\leq$ su X che massimizza L_mono_min (equivalentemente, minimizza ε^*)" è NP-difficile.

Dimostrazione. Massimizzare L_mono_min su tutti gli ordinamenti possibili equivale a trovare il cammino hamiltoniano più lungo nel grafo non orientato $G(X, \sim)$, che è NP-difficile (riduzione da Hamiltonian Path, problema NP-completo; si veda Garey e Johnson, 1979). L'identificazione segue dall'osservazione che un ordinamento totale $\leq$ che produce $L_mono_min = n - 1$ (il massimo possibile) corrisponde esattamente a un cammino hamiltoniano nel grafo: una permutazione dei vertici tale che vertici consecutivi nella permutazione siano adiacenti nel grafo. $\square$

La NP-difficoltà dell'ottimizzazione sull'ordinamento ha implicazioni importanti per l'analisi comparativa dei predicati: in generale, non è computazionalmente trattabile trovare l'ordinamento "ottimale" (quello che rende il predicato meno vago), tranne che per classi speciali di grafi (grafi intervallari, cactus, alberi) per cui esistono algoritmi polinomiali.

Capitolo 6. Cicli di Tolleranza e Ostruzioni Omologiche

6.1 Il paradosso ciclico e la sua struttura

Il Teorema 1 si occupa del caso in cui il grafo di tolleranza è connesso e si cerca di distinguere 2 estremi secondo un ordinamento lineare. Esiste però una forma diversa di paradosso di Sorite — il paradosso ciclico — che emerge quando la struttura del dominio del predicato ha natura circolare o quando il grafo di tolleranza contiene cicli. In questi casi, il problema non è di distinguere 2 estremi su una retta, ma di trovare un assegnamento coerente non costante in presenza di cicli di indistinguibilità.

Definizione 14 (Predicato ciclicamente vago). Uno spazio di tolleranza $(X, \sim)$ è ciclicamente vago se il grafo $G(X, \sim)$ contiene un ciclo indotto di lunghezza ≥ 3 , ovvero se esistono $x_1, x_2, \dots, x_k \in X$ con $k \geq 3$ tali che $x_i \sim x_{i+1}$ per ogni i (modulo k) e $x_i \not\sim x_j$ per tutti gli altri $|i - j| \geq 2$ (modulo k).

La vaghezza ciclica emerge in modo naturale per predicati su domini con struttura circolare o toroidale.

Esempio 10 (Spettro politico circolare). Si consideri uno spettro politico con 6 posizioni ordinate circolarmente: $X = \{x_1, \dots, x_6\}$ dove $x_i \sim x_{i+1}$ per $i = 1, \dots, 5$ e $x_6 \sim x_1$. Il predicato "è di sinistra" su questo spazio ha la struttura del ciclo C_6 . La modellizzazione circolare riflette l'osservazione, discussa in politologia da autori come Jean-Pierre Faye nel suo *Langages totalitaires* (1972), che in certi contesti discorsivi le posizioni di estrema sinistra e di estrema destra presentano sovrapposizioni strutturali — nazionalismo economico, anti-elitismo, critica delle istituzioni liberali — tali da rendere le 2 posizioni "simili" in alcune dimensioni. Il predicato "è di sinistra" su questo spazio ha la struttura del ciclo C_6 . Non esiste assegnamento coerente non costante: qualunque classificazione che distingue alcune posizioni non può essere estesa in modo coerente all'intero ciclo.

Esempio 11 (Spettro cromatico e predicati di colore). La percezione del colore ha una struttura ciclica descritta dall'hue circle (cerchio della tinta) della teoria del colore percettivo: partendo dal rosso (0°), si procede verso l'arancio (30°), il giallo (60°), il verde (120°), il ciano (180°), il blu (240°), il magenta (300°), e si torna al rosso ($360^\circ = 0^\circ$). La JND per la tinta è dell'ordine di $1-3^\circ$ lungo l'hue circle nelle regioni di maggiore sensibilità. Il predicato "è caldo come colore", applicato a campioni dello spettro cromatico, ha struttura ciclica: rosso è caldo, arancio è caldo, giallo è caldo, verde è borderline, blu è freddo, ma il viola è di nuovo caldo, e il viola è adiacente al rosso nel cerchio. Il ciclo non ha un "inizio" naturale.

Esempio 12 (Similarità non transitiva di Wittgenstein). Wittgenstein nelle *Philosophische Untersuchungen* (§§65–67) discute la nozione di "somiglianza di famiglia" (Familienähnlichkeit) usando l'esempio dei giochi. I giochi di carte somigliano ai giochi di strategia (entrambi hanno regole, obiettivi, vincitori), i giochi di strategia somigliano ai giochi atletici (entrambi richiedono abilità, competizione), i giochi atletici somigliano ai giochi di bambini (entrambi richiedono attività fisica), e i giochi di bambini somigliano al fare le bolle di sapone (entrambi sono divertenti e privi di scopo competitivo). Questo ciclo di somiglianze connette oggetti che a prima vista sembrano molto diversi, e mostra come la relazione "è un gioco" non abbia un confine netto ma sia tenuta insieme da catene cicliche di somiglianze.

6.2 La prima ostruzione omologica

Teorema 4 (Ostruzione omologica ciclica). Sia $(X, \sim)$ uno spazio di tolleranza tale che $G(X, \sim)$ sia esattamente il ciclo C_n con $n \geq 3$. Valgono i fatti seguenti:

- (a) Non esiste nessun assegnamento coerente $f: X \rightarrow \{\text{Vero}, \text{Falso}\}$ non costante.
- (b) $H_1(S(X, \sim); \mathbb{Z}) \cong \mathbb{Z}$, con generatore il ciclo fondamentale $[C_n]$.
- (c) Per qualsiasi campo F , $H_1(S(X, \sim); F) \cong F$, e la classe generatrice è il ciclo fondamentale.
- (d) Il tipo di omotopia di $|S(X, \sim)|$ è quello di S^1 (il cerchio), e ogni ipotetico assegnamento coerente non costante corrisponderebbe a una mappa da S^1 a $\{0, 1\}$ discontinua, il che è impossibile per ragioni topologiche.

Dimostrazione di (a). $G(X, \sim) = C_n$ è connesso per $n \geq 3$ (ogni vertice è raggiungibile da ogni altro tramite il ciclo). Per la Proposizione 1, ci sono esattamente $2^1 = 2$ assegnamenti coerenti, entrambi costanti (Falso ovunque o Vero ovunque). Nessuno dei 2 è non costante. $\square$

Dimostrazione di (b). Il complesso $S(X, \sim)$ è 1-dimensionale: ha n vertici (0-simplessi) e n archi (1-simplessi), corrispondenti agli archi del ciclo C_n . Non ci sono 2-simplessi, poiché in C_n ogni vertice ha esattamente 2 vicini, e i 2 vicini di qualunque vertice non sono adiacenti tra loro (per $n \geq 4$; per $n = 3$ il ciclo $C_3 =$ triangolo ha i 3 vertici mutualmente adiacenti e quindi 1 solo 2-simplesso, ma C_3 è il semplice pieno Δ^2 , che è contraibile, non S^1). Per $n \geq 4$, il complesso è esattamente il grafo C_n visto come complesso 1-dimensionale. Si applica la formula del numero ciclomattico:

$$\text{rk}(H_1) = |E| - |V| + c = n - n + 1 = 1$$

dove $c = 1$ è il numero di componenti connesse. Dunque $H_1(S(X, \sim); \mathbb{Z}) \cong \mathbb{Z}$. $\square$

Nota su $n = 3$: Per $C_3 = K_3$, il complesso di flag è il semplice pieno Δ^2 , che è contraibile. $H_1(\Delta^2; \mathbb{Z}) = 0$. Dunque C_3 non produce ostruzione omologica in H_1 . Questo è consistente con la Proposizione 1: $G(K_3, \sim)$ è connesso, ma il complesso di Sorite $S(K_3, \sim) = \Delta^2$ è contraibile perché la clique $\{1, 2, 3\}$ è un 2-simplesso che "riempie" il ciclo.

Dimostrazione di (c). $H_1(S^1; F) = F$ per qualunque campo F , per il teorema di omologia universale. $\square$

Dimostrazione di (d). La realizzazione geometrica $|S(X, \sim)|$ di C_n per $n \geq 4$ è omeomorfa a S^1 : n punti e n segmenti disposti in un cerchio. Una mappa continua da S^1 a $\{0, 1\}$ deve essere costante, perché $\{0, 1\}$ è un insieme finito discreto e S^1 è connesso. Dunque non esiste una mappa continua non costante da S^1 a $\{0, 1\}$, il che corrisponde all'inesistenza di assegnamenti coerenti non costanti. $\square$

6.3 Il numero di risoluzione ciclico

Per predicati ciclicamente vaghi senza ordinamento lineare, la nozione di numero di risoluzione si estende al caso "senza contrasto lineare": si cerca il minimo ε per cui esiste un assegnamento ε -coerente non costante (senza richiedere condizioni agli estremi, poiché non c'è un estremo lineare).

Definizione 15 (Numero di risoluzione ciclico). Per uno spazio di tolleranza $(X, \sim)$ senza ordinamento prefissato, il numero di risoluzione ciclico è:

$$\varepsilon^*_{\text{cicl}}(\sim) = \inf\{\varepsilon \geq 0 : \text{esiste un assegnamento } \varepsilon\text{-coerente } f: X \rightarrow [0, 1] \text{ con } \max f - \min f = 1\}$$

Teorema 5 (Formula per il numero di risoluzione ciclico). Per il ciclo C_n con $n \geq 4$ vertici e la tolleranza del ciclo (adiacenza di grado 1), si ha:

$$\varepsilon^*_{\text{cicl}}(\sim) = 1/\lfloor n/2 \rfloor$$

Dimostrazione. Si vuole trovare la funzione non costante $f: \{1, \dots, n\} \rightarrow [0, 1]$ con il minimo ε tale che $|f(i) - f(j)| \leq \varepsilon$ per ogni coppia (i, j) adiacente nel ciclo.

Parte 1: $\varepsilon^*_{\text{cicl}} \leq 1/\lfloor n/2 \rfloor$. Si costruisce esplicitamente un assegnamento $(1/\lfloor n/2 \rfloor)$ -coerente non costante. Sia $k = \lfloor n/2 \rfloor$. Si definisce la funzione "a dente di sega" simmetrica:

$$f(i) = i/k \text{ per } 1 \leq i \leq k \text{ e } f(i) = (n - i)/k \text{ per } k < i \leq n$$

(con la convenzione appropriata per i modulo n). Questa funzione raggiunge il massimo 1 in $i = k$, il minimo 0 in $i = 1$ (e $i = n$ per la parte discendente). La variazione tra elementi adiacenti è sempre $|f(i+1) - f(i)| = 1/k = 1/\lfloor n/2 \rfloor$.

Per n pari ($n = 2k$): la funzione sale da 0 a 1 in k passi ($0, 1/k, 2/k, \dots, 1$) e scende da 1 a 0 in k passi. È $(1/k)$ -coerente e non costante (raggiunge sia 0 che 1). ✓

Per n dispari ($n = 2k+1$): $\lfloor n/2 \rfloor = k$, la funzione sale in k passi e scende in $k+1$ passi (oppure in k e k). La variazione massima è ancora $1/k$.

Parte 2: $\varepsilon^*_{\text{cicl}} \geq 1/\lfloor n/2 \rfloor$. Sia f con $\max f - \min f = 1$ e ε -coerente su C_n . Siano $M = \max f$ e $m = \min f = M - 1$. Il ciclo C_n obbliga la funzione a tornare al valore iniziale dopo n passi. Il percorso "verso l'alto" da m a M percorre un numero ℓ_+ di passi, e il percorso "verso il basso" da M a m percorre $n - \ell_+$ passi.

Per il percorso ascendente (ℓ_+ passi, variazione $\leq \varepsilon$ per passo):

$$M - m \leq \ell_+ \cdot \varepsilon$$

Per il percorso discendente ($n - \ell_+$ passi, variazione $\leq \varepsilon$ per passo):

$$M - m \leq (n - \ell_+) \cdot \varepsilon$$

Sommando le 2 disuguaglianze:

$$2(M - m) \leq n \cdot \varepsilon$$

Dunque $\varepsilon \geq 2(M - m)/n$. Per f non costante, $M - m > 0$. La funzione "a dente di sega" ottimale prende $\ell_+ = \lfloor n/2 \rfloor$, il che massimizza $\min(\ell_+, n - \ell_+) = \lfloor n/2 \rfloor$. Da:

$$M - m \leq \lfloor n/2 \rfloor \cdot \varepsilon$$

si ottiene, per $M - m = 1$ (il massimo possibile per $f: X \rightarrow [0, 1]$):

$$1 \leq \lfloor n/2 \rfloor \cdot \varepsilon, \text{ ossia } \varepsilon \geq 1/\lfloor n/2 \rfloor.$$

Dunque $\varepsilon^*_{\text{cicl}} \geq 1/\lfloor n/2 \rfloor$. $\square$

La formula $\varepsilon_{\text{cicl}} = 1/\lfloor n/2 \rfloor$ mostra che cicli più lunghi hanno un numero di risoluzione ciclico più piccolo: un ciclo C_{100} ha $\varepsilon_{\text{cicl}} = 1/50$, mentre C_4 ha $\varepsilon^*_{\text{cicl}} = 1/2$. Cicli molto lunghi sono "quasi risolubili" gradatamente: anche una piccola perturbazione della coerenza permette di distinguere elementi in posizioni opposte del ciclo.

6.4 Ostruzioni omologiche per tolleranze di larghezza 2 sui cicli

Il caso della larghezza 2 sui cicli illustra come l'aumento della "portata" delle tolleranze possa eliminare le ostruzioni omologiche riempiendo i cicli con triangoli.

Proposizione 14. Per il ciclo C_n con $n \geq 6$ e la tolleranza di larghezza 2 ($i \sim j$ sse $d_{\text{circ}}(i, j) \leq 2$), il complesso di Sorite $S(C_n, \sim_2)$ ha:

- $H_1(S(C_n, \sim_2); \mathbb{Z}) \cong \mathbb{Z}$ se n non è divisibile per 3.
- $H_1(S(C_n, \sim_2); \mathbb{Z}) = 0$ se n è divisibile per 3.

Dimostrazione. Per la larghezza 2, le clique del grafo includono tutte le triple di vertici consecutivi $\{i, i+1, i+2\}$ (modulo n). Ogni tale tripla forma un triangolo (2-simplesso). Il complesso è dunque 2-dimensionale. Il numero di 2-simplessi è n (uno per ogni tripla consecutiva), il numero di 1-simplessi è $2n$ (n archi di adiacenza 1 e n archi di adiacenza 2), il numero di 0-simplessi è n .

La caratteristica di Euler è $\chi = n - 2n + n = 0$. Poiché il complesso è connesso ($H_0 \cong \mathbb{Z}$) e si ha $\chi = 1 - \text{rk}(H_1) + \text{rk}(H_2)$, si ottiene $\text{rk}(H_1) - \text{rk}(H_2) = 1$.

Per n non divisibile per 3: le n triple consecutive non possono "riempire" il ciclo senza creare un semplice vuoto; il complesso è omeomorfo a S^1 , $H_1 \cong \mathbb{Z}$, $H_2 = 0$.

Per n divisibile per 3: le triple consecutive si combinano in modo tale da creare una 2-catena che riempie il ciclo fondamentale. Il ciclo viene eliminato, $H_1 = 0$, $H_2 = 0$. Questo segue dal fatto che la somma alternata dei triangoli consecutivi (come ciclo nella catena di Borel–Moore) è il bordo del ciclo quando $n \equiv 0 \pmod{3}$. $\square$

6.5 La sequenza di Mayer–Vietoris e le ostruzioni composte

Per complessi di Sorite con struttura più ricca — non semplici cicli ma combinazioni di strutture cicliche e lineari — il calcolo dell'omologia si esegue mediante la sequenza esatta di Mayer–Vietoris, uno strumento fondamentale della topologia algebrica.

Proposizione 15 (Mayer–Vietoris per complessi di Sorite). Siano A e B due sottocomplessi di $S(X, \sim)$ con $A \cup B = S(X, \sim)$ (ogni semplice di $S(X, \sim)$ è in A o in B). Esiste la lunga sequenza esatta:

$$\dots \rightarrow H_{k+1}(A \cup B) \rightarrow H_k(A \cap B) \rightarrow H_k(A) \oplus H_k(B) \rightarrow H_k(A \cup B) \rightarrow H_{k-1}(A \cap B) \rightarrow \dots$$

Dimostrazione. Teorema di Mayer–Vietoris per coppie di sottocomplessi di un complesso CW. Si veda Munkres (1984), Teorema 25.1, o Hatcher (2002), Sezione 2.2. $\square$

Esempio 13 (Due cicli che condividono un arco). Sia $X = \{1, 2, 3, 4, 5\}$ con il grafo formato da 2 quadrati C_4 che condividono l'arco $\{3, 4\}$: il quadrato 1-2-3-4 e il quadrato 3-4-5-1. Sia A il sottocomplesso corrispondente al primo quadrato e B quello corrispondente al secondo. $A \cap B$ contiene i vertici $\{3, 4\}$ e l'arco $\{3, 4\}$.

$H_k(A)$: A è il ciclo C_4 , dunque $H_0(A) \cong \mathbb{Z}$, $H_1(A) \cong \mathbb{Z}$. $H_k(B)$: B è il ciclo C_4 , dunque $H_0(B) \cong \mathbb{Z}$, $H_1(B) \cong \mathbb{Z}$. $H_k(A \cap B)$: $A \cap B$ è un arco (1-simplesso), dunque $H_0(A \cap B) \cong \mathbb{Z}$, $H_k(A \cap B) = 0$ per $k \geq 1$.

La sequenza di Mayer–Vietoris in grado 1:

$$0 \rightarrow H_1(A) \oplus H_1(B) \rightarrow H_1(A \cup B) \rightarrow H_0(A \cap B) \rightarrow H_0(A) \oplus H_0(B)$$

$$\text{ovvero: } 0 \rightarrow \mathbb{Z} \oplus \mathbb{Z} \rightarrow H_1(A \cup B) \rightarrow \mathbb{Z} \rightarrow \mathbb{Z} \oplus \mathbb{Z}$$

La mappa $H_0(A \cap B) \rightarrow H_0(A) \oplus H_0(B)$ manda il generatore del punto componente in (generatore di A, generatore di B), dunque è iniettiva (il kernel è 0). Dalla sequenza esatta:

$$H_1(A \cup B) \cong \mathbb{Z} \oplus \mathbb{Z}$$

Ci sono 2 cicli indipendenti: i 2 quadrati originali. Questo è consistente con il numero ciclomático: $|E| - |V| + 1 = 7 - 5 + 1 = 3$... correggiamo: i vertici sono 1, 2, 3, 4, 5 (5 vertici), gli archi sono $\{1,2\}, \{2,3\}, \{3,4\}, \{4,1\}, \{4,5\}, \{5,1\}$ + l'arco $\{3,4\}$ già contato, dunque i vertici sono 5, gli archi sono 6 (non 7). Il numero ciclomático è $6 - 5 + 1 = 2$. ✓

6.6 La formula di Euler per il complesso di Sorite

Proposizione 16 (Caratteristica di Euler). Sia $S(X, \sim)$ il complesso di Sorite con f_k il numero di k -simplessi. La caratteristica di Euler è:

$$\chi(S(X, \sim)) = \sum_{k \geq 0} (-1)^k \cdot f_k = \sum_{k \geq 0} (-1)^k \cdot \text{rk}(H_k(S(X, \sim); \mathbb{Z}))$$

(il secondo membro vale quando i gruppi di omologia sono tutti liberi, cioè senza torsione)

Dimostrazione. Formula standard per complessi CW con omologia libera: la caratteristica di Euler si può calcolare sia in termini di f -vettore (combinatorio) sia in termini di numeri di Betti (topologico), e le 2 espressioni coincidono per la formula di Euler-Poincaré. Si veda Munkres (1984). □

Per il grafo di tolleranza $G(X, \sim)$ visto come complesso 1-dimensionale: $f_0 = n$, $f_1 = |E|$, $f_k = 0$ per $k \geq 2$. Dunque $\chi = n - |E|$. Poiché $\text{rk}(H_0) = c$ (numero di componenti connesse) e $\text{rk}(H_1) = \mu = |E| - n + c$ (numero ciclomático), si verifica:

$$\chi = c - \mu = c - (|E| - n + c) = n - |E| \quad \checkmark$$

La formula di Euler fornisce dunque una relazione immediata tra la struttura combinatoria del grafo e la struttura omologica del complesso. In particolare, per un grafo connesso ($c = 1$) con n vertici e m archi, $\text{rk}(H_1) = m - n + 1$.

Capitolo 7. Predicati Multidimensionali

I capitoli precedenti hanno trattato prevalentemente la vaghezza di predicati unidimensionali: predicati il cui dominio è dotato di un ordinamento totale naturale (la statura, il numero di grani, la lunghezza d'onda), e la cui relazione di tolleranza è definita da una singola soglia di indistinguibilità. La realtà dei predicati vaghi nel linguaggio naturale è tuttavia molto più ricca: la maggior parte dei predicati semanticamente interessanti — quelli che resistono alla riduzione a una singola scala — dipende simultaneamente da più parametri, ciascuno dei quali contribuisce in modo non indipendente alla classificazione complessiva.

Questo capitolo sviluppa la teoria degli spazi di tolleranza multidimensionali e dei complessi di Sorite corrispondenti, dimostrando che la struttura omologica di questi complessi è genuinamente più ricca di quella dei casi unidimensionali. In particolare, il rango di H_1 cresce polinomialmente con il numero di dimensioni e la granularità del dominio, producendo un numero molto elevato di ostruzioni omologiche indipendenti agli assegnamenti coerenti. Il numero di risoluzione ε^* assume una forma più complessa, dipendendo dalla geometria dello spazio di cammini monotoni in un DAG multidimensionale, e la formula esplicita richiede la soluzione di un problema di programmazione dinamica su un grafo a griglia.

7.1 La tolleranza prodotto e la sua geometria

La costruzione più naturale di uno spazio di tolleranza multidimensionale è quella del prodotto: dato che ogni dimensione del predicato ha la propria struttura di indistinguibilità, lo spazio prodotto eredita una relazione di tolleranza che è il prodotto delle relazioni originali. La definizione che segue formalizza questa intuizione.

Definizione 15 (Tolleranza prodotto).

Siano $(X, \sim_h)$ e $(Y, \sim_\gamma)$ due spazi di tolleranza finiti. Il prodotto cartesiano $X \times Y$ è dotato della tolleranza prodotto $\sim_{X \times Y}^5$ definita dalla seguente condizione: $(x, y) \sim_{X \times Y}^5 (x', y')$ se e solo se $(x \sim_h x' \text{ e } y = y')$ oppure $(x = x' \text{ e } y \sim_\gamma y')$. In altri termini, due punti del prodotto sono tolleranti se e solo se differiscono in esattamente una coordinata, e la variazione in quella coordinata è tollerante nello spazio originale.

L'interpretazione semantica è immediata. Il predicato 'è in forma fisicamente' dipende dalla percentuale di grasso corporeo e dalla capacità aerobica: due individui che differiscono soltanto nella percentuale di grasso corporeo di una quantità sub-JND sono indistinguibili; due individui che differiscono soltanto nella capacità aerobica di una quantità sub-JND sono indistinguibili. Ma due individui che differiscono in entrambe le dimensioni, ciascuna di una quantità sub-JND, non sono necessariamente indistinguibili: la tolleranza prodotto cattura questa struttura.

La scelta di richiedere che la variazione avvenga in esattamente una coordinata è deliberata e riflette una proprietà reale della percezione multidimensionale: la JND per variazioni combinate non è la semplice somma delle JND unidimensionali, ma dipende dalla struttura dello spazio percettivo. La tolleranza prodotto è il caso più conservativo, in cui soltanto le variazioni strettamente ortogonali producono indistinguibilità. Varianti della definizione con tolleranze 'diagonali' (dove $(x, y) \sim (x', y')$ se $x \sim_n x'$ e $y \sim_\gamma y'$ simultaneamente) producono strutture più ricche ma più difficili da analizzare; si accenna a queste varianti nella Sezione 7.6.

Proposizione 14 (Grafo di tolleranza del prodotto). Per $X = \{1, \dots, m\}$ e $Y = \{1, \dots, n\}$ con tolleranze di larghezza uniforme 1 su ciascun fattore, il grafo di tolleranza di $(X \times Y, \sim_{x^5 y})$ è la griglia $m \times n$ con adiacenza di Manhattan-1: i vertici sono le coppie (i, j) con $1 \leq i \leq m$, $1 \leq j \leq n$, e due vertici (i, j) e (i', j') sono adiacenti se e solo se $|i - i'| + |j - j'| = 1$.

Dimostrazione. La condizione di tolleranza prodotto richiede che i due punti differiscano in esattamente una coordinata con differenza 1 (per la tolleranza di larghezza 1). Questo è esattamente la condizione di adiacenza di Manhattan-1: $|i - i'| + |j - j'| = 1$ con i valori nell'insieme $\{0, 1\}$. I 4 vicini di Manhattan-1 del punto (i, j) sono $(i \pm 1, j)$ e $(i, j \pm 1)$ (quando esistono), il che coincide con la definizione di adiacenza ortogonale della griglia. $\square$

La struttura a griglia è il modello matematico per predicati con due dimensioni di variazione indipendenti. La griglia 3×3 corrisponde a un predicato con tre livelli di granularità per ciascuna delle due dimensioni. La griglia $m \times n$ con m e n grandi modella la variazione continua su entrambe le dimensioni con risoluzione finita.

Un aspetto fondamentale della tolleranza prodotto è la sua non-transitività, che eredita dalle tolleranze dei fattori. Su una griglia, la catena $(1,1) \sim (1,2) \sim (1,3) \sim (2,3) \sim (3,3)$ dimostra che $(1,1)$ e $(3,3)$, pur non essendo tolleranti tra loro, sono connessi da una catena di passi tolleranti. Questo è precisamente il paradosso di Sorite in due dimensioni: anche se ogni passo nella catena è 'semanticamente irrilevante', l'effetto accumulato attraverso entrambe le dimensioni è 'semanticamente rilevante'.

7.2 Il calcolo omologico per le griglie

Il calcolo dei gruppi di omologia del complesso di Sorite associato alla griglia è un risultato fondamentale per la teoria dei predicati multidimensionali. Il risultato principale mostra che il numero di ostruzioni omologiche indipendenti cresce rapidamente con le dimensioni della griglia.

Teorema 6 (Omologia della griglia bidimensionale). Per la griglia G_{mn} con $m \geq 2$ e $n \geq 2$, il complesso di Sorite S_{mn} ha:

$$\text{rk}(H_0(S_{mn}; \mathbb{Z})) = 1, \text{rk}(H_1(S_{mn}; \mathbb{Z})) = (m-1)(n-1), H_k(S_{mn}; \mathbb{Z}) = 0 \text{ per } k \geq 2.$$

Dimostrazione. Connessione (H_0): La griglia G_{mn} è connessa: ogni vertice (i, j) è raggiungibile dal vertice $(1,1)$ percorrendo la catena $(1,1) \sim (2,1) \sim \dots \sim (m,1) \sim (m,2) \sim \dots \sim (m,n)$. Dunque $H_0 \cong \mathbb{Z}$. Assenza di 2-simpletti: Il complesso di Sorite è il complesso di flag del grafo G_{mn} . Un 2-simplesso richiederebbe una tripla di vertici mutualmente adiacenti nel grafo. I vicini di Manhattan-1 di (i,j) sono $(i\pm 1, j)$ e $(i, j\pm 1)$. La distanza di Manhattan tra due vicini ortogonali di (i,j) è 2, non 1, dunque non sono adiacenti tra loro. Nessuna tripla è mutualmente adiacente, il complesso è 1-dimensionale e $H_k = 0$ per $k \geq 2$. Calcolo di H_1 : Poiché il complesso è 1-dimensionale, il numero ciclomático $\mu(G_{mn}) = |E| - |V| + 1 = \text{rk}(H_1)$. Per la griglia: $|V| = mn$, $|E| = m(n-1) + (m-1)n = 2mn - m - n$ (archi orizzontali più verticali). Dunque $\text{rk}(H_1) = (2mn - m - n) - mn + 1 = mn - m - n + 1 = (m-1)(n-1)$. $\square$

Il risultato $(m-1)(n-1)$ per $\text{rk}(H_1)$ merita un commento semantico. Ogni ciclo indipendente della griglia corrisponde a una 'faccia quadrata': una sequenza $(i,j) \rightarrow (i+1,j) \rightarrow (i+1,j+1) \rightarrow (i,j+1) \rightarrow (i,j)$. Ogni tale ciclo è una 'ostruzione locale': indica che nei quattro angoli del quadrato, la struttura di tolleranza crea un paradosso ciclico bidimensionale. I $(m-1)(n-1)$ quadrati della griglia sono tutti indipendenti (nessun quadrato è la somma di altri), e ciascuno genera un'ostruzione omologica distinta.

Confrontiamo con il caso unidimensionale: una catena $1-2-\dots-n$ ha $\mu = 0$ (nessun ciclo, nessuna ostruzione in H_1). La struttura bidimensionale aggiunge $(m-1)(n-1)$ ostruzioni. Per valori tipici $m = n = 10$, ci sono 81 ostruzioni omologiche indipendenti. Questo spiega perché i predicati genuinamente bidimensionali come 'è in forma' siano semanticamente più 'vagi' dei predicati unidimensionali nel senso omologico: la struttura del paradosso è molto più ricca.

Corollario 6.1 (Generatori di H_1). $H_1(S_{mn}; \mathbb{Z})$ è generato liberamente dai cicli fondamentali dei $(m-1)(n-1)$ quadrati elementari $\gamma_{ij} = [(i,j) \rightarrow (i+1,j) \rightarrow (i+1,j+1) \rightarrow (i,j+1) \rightarrow (i,j)]$ per $1 \leq i \leq m-1, 1 \leq j \leq n-1$. Ogni 1-ciclo del complesso è una combinazione intera di questi generatori.

Dimostrazione. I cicli γ_{ij} sono non-banali in H_1 (non sono bordi di 2-catene, poiché non esistono 2-simpletti) e linearmente indipendenti (poiché ogni arco verticale $(i,j) \rightarrow (i,j+1)$ compare in esattamente i cicli $\gamma_{i-1,j}$ e $\gamma_{i,j}$ con segni opposti, il che impedisce cancellazioni non banali). Il loro numero è $(m-1)(n-1) = \text{rk}(H_1)$, dunque formano una base. $\square$

La struttura del complesso di Sorite per le griglie ha un'elegante interpretazione combinatoria. Si consideri il dual del grafo a griglia: i vertici del dual sono le facce quadrate della griglia, e due facce sono adiacenti se condividono un arco. Il dual della griglia $m \times n$ è la griglia $(m-1) \times (n-1)$. Gli elementi di H_1 corrispondono a funzioni sui vertici del grafo duale, e la struttura di H_1 rispecchia la geometria del grafo duale. Questa corrispondenza, formalizzata dalla dualità di Poincaré per complessi cellulari, fornisce un'ulteriore prospettiva sulla struttura omologica del complesso di Sorite.

7.3 Il numero di risoluzione per i predicati prodotto

Il calcolo del numero di risoluzione ε^* per predicati multidimensionali richiede di identificare il cammino monotono minimo nell'ordinamento totale indotto sul prodotto cartesiano. La scelta dell'ordinamento è cruciale: ordinamenti diversi sul prodotto cartesiano producono numeri di risoluzione diversi.

Proposizione 15 (Numero di risoluzione bidimensionale — ordinamento lessicografico). Per la griglia G_{mn} con tolleranza di larghezza 1 e l'ordinamento lessicografico $(i,j) < (i',j')$ se e solo se $i < i'$, oppure $i = i'$ e $j < j'$, il numero di risoluzione è:

$$\varepsilon^*(\sim_x^5) = 1 / (m + n - 2)$$

Dimostrazione. Parte 1 ($\varepsilon^* \leq 1/(m+n-2)$): Il cammino monotono minimo da $(1,1)$ a (m,n) nell'ordinamento lessicografico è la sequenza che percorre prima tutta la prima riga, poi sale di riga, e così via. Formalmente: $(1,1), (1,2), \dots, (1,n), (2,n), (2,n-1), \dots$, ma questo non è monotono nell'ordinamento lessicografico perché $(2,n-1) < (2,n)$. Il cammino monotono minimo è invece: $(1,1), (1,2), \dots, (1,n), (2,n), (3,n), \dots, (m,n)$. La lunghezza è $(n-1) + (m-1) = m+n-2$. La costruzione della Parte 1 del Teorema 3 dà un assegnamento $(1/(m+n-2))$ -coerente. Parte 2 ($\varepsilon^* \geq 1/(m+n-2)$): Ogni assegnamento ε -coerente monotono con le condizioni agli estremi deve assegnare 0 a $(1,1)$ e 1 a (m,n) . La somma telescopica lungo il cammino di lunghezza $m+n-2$ dà $1 \leq (m+n-2)\varepsilon$, dunque $\varepsilon \geq 1/(m+n-2)$. $\square$

La formula $\varepsilon^* = 1/(m+n-2)$ ha un'interpretazione geometrica chiara: il cammino monotono minimo in una griglia $m \times n$ deve percorrere esattamente $m-1$ passi verso il basso e $n-1$ passi verso destra (o in qualunque combinazione monotona), per un totale di $m+n-2$ passi. La lunghezza del cammino non dipende dall'ordinamento totale scelto (lessicografico o altro), purché l'ordinamento sia compatibile con la struttura della griglia nel senso che rende monotoni i percorsi dalla cella in basso a sinistra a quella in alto a destra.

Proposizione 15.1 (Confronto con il caso unidimensionale). Per una catena lineare di n punti con tolleranza uniforme 1, $\varepsilon^* = 1/(n-1)$. Per la griglia $m \times n$, $\varepsilon^* = 1/(m+n-2)$. Poiché $m+n-2 \geq \max(m-1, n-1)$ per $m, n \geq 2$, il predicato bidimensionale è meno vago (in senso graduato) del peggiore dei due predicati unidimensionali. Il numero totale di punti della griglia è mn , ma la

lunghezza del cammino minimo è solo $m+n-2$, che è molto più piccola di $mn-1$ (la lunghezza del cammino hamiltoniano). Questo mostra che la struttura bidimensionale permette 'scorciatoie' nel cammino monotono che riducono ε^* rispetto al caso unidimensionale con lo stesso numero di punti.

7.4 La formula di Künneth per predicati prodotto

La formula di Künneth è uno degli strumenti fondamentali dell'algebra omologica per calcolare l'omologia di spazi prodotto. La sua applicazione al contesto degli spazi di tolleranza richiede una distinzione importante tra il prodotto topologico delle realizzazioni geometriche e il complesso di Sorite del prodotto degli spazi di tolleranza.

Teorema 7 (Künneth per realizzazioni geometriche). Siano $S_X = S(X, \sim_X)$ e $S_Y = S(Y, \sim_Y)$ due complessi simpliciali finiti con omologia libera (senza torsione). Per il prodotto topologico $|S_X| \times |S_Y|$, vale:

$$H_k(|S_X| \times |S_Y|; \mathbb{Z}) \cong \bigoplus_{i+j=k} H_i(S_X; \mathbb{Z}) \otimes H_j(S_Y; \mathbb{Z})$$

Dimostrazione. Teorema di Künneth standard per la categoria dei CW-complessi con omologia libera. Si veda Munkres (1984), Capitolo 7, o Hatcher (2002), Sezione 3.2. La formula si semplifica in assenza di torsione poiché il termine correttivo di Tor scompare. $\square$

Osservazione 3 (Distinzione cruciale tra prodotto topologico e prodotto di tolleranza). Il Teorema 7 si applica al prodotto topologico delle realizzazioni geometriche, non al complesso di Sorite del prodotto degli spazi di tolleranza. I due oggetti sono in generale diversi. Il complesso di Sorite $S(X \times Y, \sim_{X \times Y})$ è tipicamente molto più piccolo del prodotto topologico $|S_X| \times |S_Y|$ in termini di dimensione e complessità topologica.

Per illustrare la differenza, si confrontino i risultati per la griglia 3×3 . Il prodotto topologico delle realizzazioni geometriche delle due catene $1-2-3$ è il quadrato $[0,1] \times [0,1]$, che ha $H_0 = \mathbb{Z}$, $H_1 = 0$, $H_2 = 0$ (il quadrato è contraibile). Il complesso di Sorite della griglia 3×3 ha invece $H_1 = \mathbb{Z}^4$ (quattro cicli indipendenti). La struttura omologica del complesso di Sorite è molto più ricca di quella del prodotto topologico.

Questa osservazione ha implicazioni importanti: non è possibile calcolare il sorites barcode di un predicato multidimensionale semplicemente combinando i barcode dei predicati unidimensionali componenti. La struttura bidimensionale del grafo di tolleranza produce topologia che non è visibile nelle singole dimensioni. Questa è una ragione fondamentale per cui la teoria topologica è necessaria per i predicati multidimensionali, e non è sostituibile da una semplice aggregazione dei risultati unidimensionali.

7.5 Il rango di H_1 per griglie d-dimensionali

La generalizzazione a dimensioni superiori è diretta. Una griglia d-dimensionale di formato $n_1 \times n_2 \times \dots \times n_d$ con tolleranza di larghezza 1 su ciascuna direzione produce un complesso simpliciale la cui omologia si calcola combinatorialmente.

Proposizione 16 (Rango di H_1 per griglie d-dimensionali). Per la griglia d-dimensionale $n_1 \times \dots \times n_d$ con tolleranza di larghezza 1 su ciascuna direzione, il complesso di Sorite è 1-dimensionale (nessun k-simplesso per $k \geq 2$) e:

$$\text{rk}(H_1) = \sum_{i=1}^d (n_i - 1) \cdot \prod_{k \neq i} n_k - \prod_{i=1}^d n_i + 1$$

Dimostrazione. 1-dimensionalità: Il grafo di tolleranza della griglia d-dimensionale ha adiacenza di Manhattan-1. In questo grafo, due vicini qualsiasi di un vertice v differiscono da v in una coordinata ciascuno, e differiscono tra loro in due coordinate (con contributo 1 ciascuna), per una distanza di Manhattan totale di 2. Non essendo adiacenti, nessuna tripla è mutualmente adiacente, e il complesso è 1-dimensionale. Formula per $\text{rk}(H_1)$: Si calcola il numero cicломatico direttamente. Il numero di vertici è $|V| = n_1 \cdots n_d = \prod n_i$. Il numero di archi è $|E| = \sum_{i=1}^d (n_i - 1) \cdot \prod_{k \neq i} n_k$ (archi nella direzione i , moltiplicati per il numero di ‘strati’ nelle altre direzioni). La formula del numero cicломatico per il grafo connesso ($c=1$) dà direttamente $\text{rk}(H_1) = |E| - |V| + 1 = \sum_{i=1}^d (n_i - 1) \cdot \prod_{k \neq i} n_k - \prod n_i + 1$. Non vi è bisogno di sviluppare ulteriormente: questa è la formula esatta. $\square$

Per una griglia cubica d-dimensionale di lato n ($n_1 = \dots = n_d = n$), la formula diventa:

$$\text{rk}(H_1) = d(n-1)n^{d-1} - n^d + 1 = n^{d-1}((d-1)n - d) + 1$$

La formula corretta $\text{rk}(H_1) = n^{d-1}((d-1)n - d) + 1$ mostra che il numero di ostruzioni omologiche cresce come $(d-1)n^d$ per n grande. Per un predicato tridimensionale ($d=3$) con 10 punti per dimensione ($n=10$), ci sono $10^2(2 \cdot 10 - 3) + 1 = 100 \cdot 17 + 1 = 1701$ ostruzioni omologiche indipendenti. Si noti che la formula precedentemente enunciata $d(d-1)/2 \cdot (n-1)^2 \cdot n^{d-2}$ era errata: la verifica diretta sul cubo $2 \times 2 \times 2$ (8 vertici, 12 archi) dà $\text{rk}(H_1)=5$, mentre la formula corretta dà $2(1 \cdot 2 - 3) + 1 = -1$, impossibile. La formula esatta si calcola sempre come $|E| - |V| + 1$.

7.6 Vaghezza multidimensionale: struttura dello spazio di assegnamenti

Per i predicati multidimensionali, lo spazio degli assegnamenti ε -coerenti monotoni ha una struttura geometrica più ricca di quella dei predicati unidimensionali. In particolare, la degenerazione del minimo — il numero di assegnamenti ottimali — è determinata dalla struttura del DAG monotono nella griglia, che ammette molti cammini minimi di uguale lunghezza.

Proposizione 17 (Degenerazione degli assegnamenti ottimali per griglie). Per la griglia $m \times n$ con ordinamento lessicografico, il numero di cammini monotoni minimi da $(1,1)$ a (m,n) è il coefficiente binomiale $\beta(m+n-2, m-1)$.

Dimostrazione. Un cammino monotono minimo consiste di $m-1$ passi 'verso il basso' (incremento della prima coordinata) e $n-1$ passi 'verso destra' (incremento della seconda coordinata), in qualche ordine. Il numero di sequenze di $m+n-2$ passi con $m-1$ di un tipo e $n-1$ dell'altro è il coefficiente binomiale $\beta(m+n-2, m-1)$. $\square$

Il coefficiente binomiale $\beta(m+n-2, m-1)$ cresce esponenzialmente in m e n (è dell'ordine di $2^{\sqrt{m+n}}$ per $m \approx n$). Questo significa che ci sono esponenzialmente molti assegnamenti ottimali: la struttura di vaghezza del predicato non determina univocamente la funzione di membership 'corretta', ma soltanto la classe di equivalenza di funzioni di membership ottimali. Questo è un argomento matematico contro l'esistenza di una funzione di membership 'canonica' per i predicati multidimensionali, e fornisce una risposta parziale alla critica delle logiche fuzzy che lamenta la difficoltà di scegliere la funzione di membership.

La struttura degli assegnamenti ottimali per griglie ha anche un'interpretazione filosofica: ci sono $(m-1)(n-1)$ gradi di libertà omologici (le ostruzioni in H_1) e $\beta(m+n-2, m-1)$ cammini ottimali. La corrispondenza tra questi due invarianti è un tema che merita sviluppo futuro: si congettura che il gruppo fondamentale dello spazio degli assegnamenti ottimali sia direttamente correlato con H_1 del complesso di Sorite.

Capitolo 8. La Teoria Spettrale degli Spazi di Tolleranza

La teoria spettrale dei grafi fornisce strumenti quantitativi potenti per l'analisi della struttura di connessione e delle proprietà di mescolamento dei grafi di tolleranza. In questo capitolo si sviluppa la connessione tra gli autovalori del laplaciano normalizzato del grafo di tolleranza e le proprietà di vaghezza del predicato associato. Il risultato principale è che il gap spettrale — la differenza tra il primo e il secondo autovalore — misura in modo quantitativo la 'quasi-vaghezza' del predicato: quanto il grafo di tolleranza si avvicina alla disconnessione senza raggiungerla.

La teoria spettrale degli spazi di tolleranza è rilevante non soltanto per la comprensione matematica della vaghezza, ma anche per le applicazioni computazionali: gli autovalori del laplaciano si calcolano efficientemente, e forniscono quindi indicatori computabili della struttura di vaghezza che possono essere estratti da dati empirici senza richiedere il calcolo completo del sorites barcode.

8.1 L'operatore di diffusione e le sue proprietà

Definizione 16 (Operatore di diffusione).

Per uno spazio di tolleranza finito $(X, \sim)$ con grafo di tolleranza

$G(X, \sim)$, l'operatore di diffusione $T_{\sim}: \mathbb{R}^X \rightarrow \mathbb{R}^X$ è definito da:

$$(T_{\sim} f)(x) = \sum_{y: y \sim x, y \neq x} f(y) / \deg(x)$$

dove $\deg(x) = |\{y \in X : y \sim x, y \neq x\}|$ è il grado di x nel grafo.

L'operatore $T_{\sim}$ coincide con la matrice della passeggiata casuale normalizzata sul grafo $G(X, \sim)$.

L'interpretazione semantica dell'operatore di diffusione è quella di un 'processo di consenso': se ogni elemento del dominio del predicato aggiorna il proprio valore di verità alla media dei valori di verità dei suoi vicini tolleranti, l'iterazione di questo processo converge a un assegnamento che è costante su ogni componente connessa. La velocità di questa convergenza è determinata dagli autovalori di $T_{\sim}$.

Proposizione 17 (Proprietà spettrali di $T_{\sim}$). Valgono i seguenti fatti:

- (a) $T_{\sim}$ è un operatore lineare con $\|T_{\sim}\|_{\infty} \leq 1$.
- (b) $T_{\sim}$ è simmetrico rispetto al prodotto interno pesato $\langle f, g \rangle_{\pi} = \sum_x \pi(x) f(x) g(x)$ dove $\pi(x) = \deg(x) / (2|E|)$ è la distribuzione stazionaria della passeggiata casuale.
- (c) Gli autovalori di $T_{\sim}$ sono tutti nell'intervallo $[-1, 1]$.
- (d) La molteplicità dell'autovalore $\lambda = 1$ è uguale al numero di componenti connesse k di $G(X, \sim)$.
- (e) $\lambda = -1$ è autovalore se e solo se $G(X, \sim)$ contiene una componente bipartita; la sua molteplicità è il numero di tali componenti.

Dimostrazione. (a): Per ogni $f \in \mathbb{R}^X$ e $x \in X$, $|(T_{\sim} f)(x)| = |\sum_{y \sim x} f(y) / \deg(x)| \leq \sum_{y \sim x} |f(y)| / \deg(x) \leq \|f\|_{\infty}$. (b): Si verifica che $\langle T_{\sim} f, g \rangle_{\pi} = \langle f, T_{\sim} g \rangle_{\pi}$ usando la simmetria della relazione $\sim$ e la definizione di π . (c)-(d): La simmetria implica che gli autovalori sono reali. Il bound $|\lambda| \leq 1$ segue da (a) e dal fatto che $T_{\sim}$ è stocastica (righe che sommano a 1). Gli autovettori di autovalore 1 sono esattamente le funzioni costanti su

ciascuna componente connessa. (e): Teoria classica delle passeggiate casuali su grafi. Si veda Levin, Peres e Wilmer (2009), Capitolo 12. $\square$

Corollario 8.1 (Il gap spettrale come misura di quasi-vaghezza). Il gap spettrale $1 - \lambda_2$ (dove λ_2 è il secondo autovalore più grande di $T_{\sim}$) misura quantitativamente la 'quasi-connessione' del grafo di tolleranza. Un gap spettrale piccolo ($1 - \lambda_2 \approx 0$) corrisponde a un grafo che è quasi-disconnesso — la passeggiata casuale rimane intrappolata per lungo tempo in una parte del grafo — il che semanticamente corrisponde a un predicato con una zona di confine pronunciata.

Il legame tra il gap spettrale e la struttura di vaghezza è il seguente. Se $G(X, \sim)$ è connesso ma 'quasi-disconnesso' nel senso che esiste un sottoinsieme S tale che il numero di archi che attraversano il confine di S è molto piccolo rispetto al volume di S , allora λ_2 è vicino a 1 e il gap è piccolo. In termini semantici, il confine di S corrisponde a una 'zona grigia' del predicato: gli elementi in S sono 'quasi' classificabili come Vero (o come Falso) senza violare significativamente la coerenza, ma qualunque classificazione netta richiederebbe di 'tagliare' alcuni archi di tolleranza nella zona di confine.

Il gap spettrale non è lo stesso dell'inverso del numero di risoluzione ε^* . Le due quantità misurano aspetti diversi della vaghezza: ε^* misura il costo minimo globale di classificazione (quanto si deve ammorbidire la coerenza per permettere una distinzione degli estremi), mentre il gap spettrale misura la struttura locale del collo di bottiglia (dove si concentra la difficoltà di classificazione nel dominio). Per predicati con struttura di tolleranza uniforme, le due quantità sono inversamente proporzionali; per predicati con struttura non uniforme, possono discostarsi significativamente.

8.2 Il laplaciano normalizzato e la disuguaglianza di Cheeger

Definizione 17 (Laplaciano normalizzato).

Il laplaciano normalizzato del grafo $G(X, \sim)$ è la matrice $L = I - T_{\sim}$, con: $L[x,y] = 1$ se $x = y$

$$L[x,y] = -1/\deg(x) \text{ se } x \sim y, x \neq y$$

$L[x,y] = 0$ altrimenti.

L è simmetrico semidefinito positivo, con autovalori $0 = \mu_0 \leq \mu_1 \leq \dots$

$\leq \mu_{\{n-1\}}$. Il minimo autovalore non nullo μ_1 è chiamato costante di

Fiedler o connettività algebrica del grafo.

Teorema 8 (Disuguaglianza di Cheeger discreta, Alon-Milman 1985). Per il grafo $G(X, \sim)$ con laplaciano normalizzato e primo autovalore non nullo μ_1 , si ha:

$$h^2/2 \leq \mu_1 \leq 2h$$

dove h è la costante di Cheeger (isoperimetrica discreta):

$$h = \min_{\{S \subset X, 0 < |S| \leq n/2\}} |\partial S| / \min(\text{vol}(S), \text{vol}(X \setminus S))$$

e dove $|\partial S|$ è il numero di archi tra S e $X \setminus S$, $\text{vol}(S) = \sum_{x \in S} \deg(x)$.

Dimostrazione. Alon e Milman (1985), 'Lambda_1, Isoperimetric Inequalities for Graphs, and Superconcentrators'. Il bound superiore $h \leq \sqrt{2\mu_1}$ segue dall'argomento di Cheeger standard applicato al vettore di Fiedler: la bipartizione indotta dal vettore di Fiedler ha costante di Cheeger al più $2\mu_1$. Il bound inferiore $\mu_1 \leq 2h$ segue dall'analisi di Rayleigh del quoziente per il vettore indicatore di S . $\square$

La costante di Cheeger h misura quanto il grafo di tolleranza si avvicina alla disconnessione: $h = 0$ se e soltanto se il grafo è disconnesso (o ha una componente bipartita con il bordo vuoto). Un valore piccolo di h indica l'esistenza di un 'collo di bottiglia': un sottoinsieme S tale che il confine $|\partial S|$ è piccolo rispetto al volume $\text{vol}(S)$. In termini semantici, questo corrisponde a un predicato che ha una 'zona di confine' relativamente netta, nonostante non abbia un confine esatto.

Esempio. Per il predicato 'è alto' con stature distribuite uniformemente tra 150 e 200 cm e tolleranza di 5 cm, il grafo di tolleranza è una catena di 51 nodi. Il sottoinsieme $S = \{150, \dots, 175\}$ ha $|S| = 26$, $|\partial S| = 1$ (l'arco tra 175 e 176 cm), $\text{vol}(S) = 2 \cdot (26-1) + 2 = 52$ (i nodi interni hanno grado 2, quelli agli estremi grado 1). Dunque $h \leq 1/52 \approx 0.019$. Il gap spettrale $\mu_1 \leq 2h \approx 0.038$, indicando una struttura quasi-disconnessa con un collo di bottiglia intorno ai 175 cm. Questo modella l'intuizione che 175 cm sia una 'soglia psicologica' per la classificazione come alto.

La disuguaglianza di Cheeger ha anche importanti conseguenze computazionali: fornisce un algoritmo per trovare approssimativamente il collo di bottiglia del grafo di tolleranza (e quindi la zona di confine del predicato) calcolando il vettore di Fiedler. Algoritmi efficienti per il calcolo del secondo autovalore del laplaciano (basati su metodi di potenza o decomposizioni spettrali) permettono di trovare questa struttura in tempo $O(|E| \cdot \text{polilog}(n))$ per grafi sparsi.

8.3 Convergenza della passeggiata casuale e struttura del confine

Teorema 9 (Convergenza della passeggiata casuale). Per il grafo di tolleranza $G(X, \sim)$ connesso e non bipartito, la passeggiata casuale con matrice di transizione $T_{\sim}$ converge alla distribuzione stazionaria π (dove $\pi(x) = \deg(x)/(2|E|)$). Il tempo di mescolamento t_{mix} — il minimo t tale che la distanza in variazione totale dalla distribuzione stazionaria sia $\leq 1/4$ — soddisfa:

$$t_{\text{mix}} \leq (1/\mu_1) \cdot \log(1/(\pi_{\min} \cdot \varepsilon_{\text{conv}}))$$

dove $\pi_{\min} = \min_x \pi(x)$ e $\varepsilon_{\text{conv}}$ è la precisione desiderata. (Si veda Levin, Peres e Wilmer, 2009.)

Il tempo di mescolamento t_{mix} misura quanto velocemente la passeggiata casuale 'dimentica' la posizione iniziale e si distribuisce uniformemente (rispetto alla distribuzione stazionaria) sul grafo. In termini di vaghezza, il tempo di mescolamento misura la velocità con cui l'informazione 'si diffonde' attraverso le catene di tolleranza.

Se t_{mix} è grande (gap spettrale piccolo), la passeggiata rimane 'intrappolata' nella regione iniziale per lungo tempo: questo corrisponde a un predicato con una netta separazione tra la zona dei casi chiaramente Vero e quella dei casi chiaramente Falso, con una transizione lenta nella zona di confine. Se t_{mix} è piccolo (gap spettrale grande), la passeggiata si mescola rapidamente: questo corrisponde a un predicato in cui la struttura di tolleranza è 'isotropa' e non c'è una zona di confine particolarmente concentrata.

Proposizione 18 (Relazione tra t_{mix} e ε^*). Per grafi di tolleranza con struttura uniforme (grafi regolari o quasi-regolari), il tempo di mescolamento e il numero di risoluzione sono approssimativamente inversi l'uno dell'altro: $t_{\text{mix}} \approx 1/\varepsilon^*$ (a meno di fattori logaritmici dipendenti dal grado). Questa relazione suggerisce che il 'tempo caratteristico' del paradosso di Sorite (il numero di passi nella catena di indistinguibilità da un estremo all'altro) è dello stesso ordine del tempo di mescolamento.

Dimostrazione. La dimostrazione usa il fatto che per grafi regolari di grado k , $\pi(x) = 1/n$ per ogni x , e $\mu_1 \approx c/t_{\text{mix}}$ per una costante c che dipende dal grado. Il numero di risoluzione $\varepsilon^* = 1/L_{\text{mono_min}} \approx k/(n-1)$ per tolleranze uniformi. La relazione $t_{\text{mix}} \approx n/(ck) \approx 1/(c\varepsilon^*)$ segue da un argomento di scaling. $\square$

8.4 Il vettore di Fiedler e la bipartizione ottimale del dominio

Teorema 10 (Vettore di Fiedler e bipartizione approssimata). L'autovettore del laplaciano normalizzato corrispondente al minimo autovalore non nullo μ_1 — il vettore di Fiedler $v_1 \in \mathbb{R}^X$ — fornisce una bipartizione approssimata del dominio del predicato: i vertici con $v_1(x) > 0$ tendono al 'Vero' e quelli con $v_1(x) < 0$ tendono al 'Falso'. La qualità di questa bipartizione è garantita dalla disuguaglianza di Cheeger: la costante di Cheeger della bipartizione è al più $2\sqrt{2\mu_1}$.

Dimostrazione. Il vettore di Fiedler minimizza il quoziente di Rayleigh $\langle v, Lv \rangle_\pi / \langle v, v \rangle_\pi$ tra tutti i vettori ortogonali alla distribuzione stazionaria π . La bipartizione indotta dalla soglia zero del vettore di Fiedler è approssimativamente ottimale per la costante di Cheeger, per il teorema di Cheeger discreto di Alon e Milman (1985). $\square$

Applicazione alla vaghezza. Il vettore di Fiedler applicato al grafo di tolleranza di un predicato vago fornisce la migliore approssimazione a un assegnamento binario coerente: anziché richiedere coerenza esatta (impossibile quando il grafo è connesso), si cerca l'assegnamento che massimizza la 'quasi-coerenza' in senso spettrale. Il vettore di Fiedler, normalizzato in $[0,1]$, è una candidata naturale per la funzione di membership di una logica fuzzy derivata dalla struttura di tolleranza.

Il confronto tra il vettore di Fiedler e la funzione di membership ottimale del Teorema 3 rivela una distinzione concettuale importante. La funzione di membership ottimale (quella che realizza ε^*) è ottimale nel senso globale: minimizza il massimo scarto tra elementi tolleranti sul cammino monotono minimo. Il vettore di Fiedler è ottimale nel senso spettrale: minimizza la variazione quadratica media tra elementi tolleranti. Le due ottimalità producono funzioni diverse, che corrispondono a due diverse formalizzazioni dell'intuizione 'la funzione di membership deve rispettare la struttura di tolleranza'.

Esempio. Per il predicato 'è calvo' su una catena lineare di 101 valori (numero di capelli da 0 a 100.000, con 1000 per valore), il vettore di Fiedler produce una funzione sigmoide smussata (simile alla distribuzione logistica) centrata nella zona di confine, mentre la funzione ottimale del Teorema 3 produce una funzione lineare a tratti con ginocchia ai nodi del cammino monotono minimo. La funzione sigmoide è più 'biologicamente plausibile' (rispecchia le curve psicometriche osservate), ma la funzione lineare a tratti è ottimale nel senso del numero di risoluzione.

8.5 Passeggiate casuali, vaghezza e processi di categorizzazione

Un collegamento più profondo tra la teoria spettrale e la semantica cognitiva emerge considerando la passeggiata casuale come modello del processo di categorizzazione. Quando un individuo deve classificare un oggetto rispetto a un predicato vago, il processo cognitivo può essere modellato come una passeggiata casuale nello spazio delle istanze: si parte dall'oggetto da classificare e si 'esplorano' istanze simili, aggiornando la classificazione in base alle classificazioni degli oggetti esplorati.

In questo modello, la distribuzione stazionaria π della passeggiata casuale corrisponde alla 'distribuzione di attenzione' della comunità linguistica: gli oggetti con grado alto (molti vicini tolleranti) ricevono più attenzione e sono più centrali nella struttura di categorizzazione. Il tempo di mescolamento t_{mix} corrisponde al tempo cognitivo necessario per stabilizzare la classificazione in una zona di confine.

Questa interpretazione suggerisce un'ipotesi empiricamente testabile: il tempo di reazione nei compiti di categorizzazione dovrebbe essere correlato con il tempo di mescolamento locale del grafo di tolleranza nella zona dell'oggetto da classificare. Gli oggetti in zone del grafo con bassa costante di Cheeger locale (colli di bottiglia) dovrebbero produrre tempi di reazione più lunghi. Esperimenti di psicologia cognitiva sul tempo di risposta per oggetti 'borderline' (Osherson e Smith, 1981; Hampton, 1998) sono consistenti con questa previsione, anche se una verifica quantitativa rigorosa richiederebbe dati molto più fini di quelli attualmente disponibili.

Proposizione 19 (Spettro delle griglie e vaghezza bidimensionale). Per la griglia $m \times n$, gli autovalori del laplaciano normalizzato sono:

$$\mu_{ij} = 1 - (\cos(\pi i/(m-1)) + \cos(\pi j/(n-1))) / 2 \text{ per } 0 \leq i \leq m-1, 0 \leq j \leq n-1$$

e il primo autovalore non nullo (gap spettrale) è $\mu_1 = (1 - \cos(\pi/\max(m-1, n-1))) / 2 \approx \pi^2/(2 \cdot \max(m-1, n-1)^2)$ per m, n grandi. [Verifica: per $m=n=2$ (griglia 2×2 = ciclo C_4), $\mu_1 = (1 - \cos(\pi/1))/2 = (1+1)/2 = 1$, che coincide con gli autovalori noti di C_4 : $\{0, 1, 1, 2\}$. $\checkmark$]

Dimostrazione. Gli autovalori del laplaciano normalizzato della griglia si calcolano per separazione di variabili: $L = L_m \otimes I_n + I_m \otimes L_n$, dove L_m e L_n sono i laplaciani normalizzati delle catene di m e n nodi rispettivamente. Gli autovalori del laplaciano normalizzato della catena P_n sono $1 - \cos(\pi k/(n-1))$ per $k = 0, \dots, n-1$ (con la giusta normalizzazione ai vertici di grado 1). Gli autovalori di L sono dunque le somme $\mu_{ij} = \lambda_i^{(m)} + \lambda_j^{(n)} = (1 - \cos(\pi i/(m-1))) + (1 - \cos(\pi j/(n-1))) - 1 = 1 - (\cos(\pi i/(m-1)) + \cos(\pi j/(n-1))) / 2$ per la normalizzazione adeguata. La formula verifica $\mu_{00}=0$ e $\mu_{m-1, n-1}=2$, in accordo con la semidefinita positività e il bound superiore 2 del laplaciano normalizzato. $\square$

La formula per μ_1 mostra che il gap spettrale della griglia $m \times n$ è dell'ordine di $1/\max(m-1, n-1)^2$, mentre per una catena di $m \cdot n$ nodi (con lo stesso numero totale di elementi) sarebbe dell'ordine di $1/(m \cdot n)^2$. Il gap spettrale della griglia è molto più grande: la struttura bidimensionale 'riduce' il tempo di mescolamento rispetto alla catena unidimensionale equivalente. Questo è un altro modo in cui i predicati bidimensionali sono meno vaghi dei predicati unidimensionali con lo stesso numero di istanze.

Capitolo 9. Omologia Persistente e il Sorites Barcode

L'omologia persistente è uno strumento della topologia algebrica computazionale che studia come la struttura omologica di una famiglia di spazi topologici (filtrazione) evolve al variare di un parametro. Applicata alla filtrazione di Sorite — la famiglia di complessi di Sorite parametrizzata dal raggio di tolleranza — l'omologia persistente produce il sorites barcode: una rappresentazione compatta e stabile della struttura di vaghezza di un predicato. Questo capitolo sviluppa questa costruzione in modo sistematico, dimostrando la stabilità del barcode rispetto a perturbazioni della metrica e la relazione tra il barcode e gli altri invarianti (numero di risoluzione, costante di Cheeger) introdotti nei capitoli precedenti.

Il sorites barcode rappresenta una 'firma topologica' del predicato: due predicati con la stessa struttura di indistinguibilità (anche se definiti su domini diversi) hanno lo stesso barcode, e la distanza bottleneck tra i barcode di due predicati misura quanto diversa è la loro struttura di vaghezza. Questa proprietà di stabilità è essenziale per le applicazioni: permette di confrontare predicati estratti da dati empirici diversi e di tracciare l'evoluzione della struttura di vaghezza nel tempo.

9.1 La filtrazione di Sorite

Definizione 18 (Filtrazione di Sorite).

Per uno spazio metrico finito (X, d) e $r \geq 0$, si definisce la
relazione di tolleranza $\sim_r$: $x \sim_r y$ se e solo se $d(x, y) \leq r$. La

filtrazione di Sorite è la famiglia di complessi simpliciali:

$$\{S(X, \sim_r)\}_{r \geq 0}$$

crescente nel senso che $r \leq s$ implica $S(X, \sim_r) \subseteq S(X, \sim_s)$: ogni

simpleso che esiste al raggio r esiste anche al raggio s .

La filtrazione di Sorite descrive come la struttura di indistinguibilità del predicato si arricchisce al crescere della soglia di tolleranza. Per $r = 0$, il complesso è discreto: nessun due elementi sono tolleranti (a meno di elementi coincidenti). Per $r \rightarrow \infty$, il complesso tende al semplice pieno su X : tutti gli elementi sono tolleranti tra loro. Il passaggio dal discreto al semplice pieno avviene attraverso una serie di 'eventi topologici': la comparsa di nuovi archi (quando la distanza tra due elementi supera la soglia r), la formazione di nuovi triangoli (quando tre elementi mutuamente indistinguibili compaiono insieme), e così via.

Osservazione 4 (Identificazione con Vietoris-Rips).

La filtrazione di Sorite è identica alla filtrazione di Vietoris-Rips $VR(X, r)$: per la definizione del complesso di Sorite come complesso di flag del grafo di tolleranza, $S(X, \sim_r)$ coincide con $VR(X, r)$ per ogni r . Tutti i risultati della TDA sviluppati per la filtrazione di Vietoris-Rips si applicano direttamente alla filtrazione di Sorite.

Questa identificazione è importante perché permette di usare direttamente i software di calcolo dell'omologia persistente sviluppati per la TDA (Gudhi, Ripser, Eirene) per calcolare il sorites barcode di un predicato linguistico: basta fornire la matrice delle distanze tra le istanze del predicato. La complessità computazionale di questo calcolo è descritta nel Capitolo 13.

La filtrazione di Sorite ha una proprietà aggiuntiva rispetto a una generica filtrazione di Vietoris-Rips: l'insieme X è dotato di un ordinamento totale (l'ordinamento del predicato, es. dal meno alto al più alto per 'è alto'), e questo ordinamento permette di definire il sorites barcode arricchito descritto nella Definizione 21.

9.2 Il diagramma di persistenza

Definizione 19 (Diagramma di persistenza).

Il diagramma di persistenza PD_k della filtrazione di Sorite è il

multiinsieme di punti $(r_{\text{birth}}, r_{\text{death}}) \in [0, \infty]^2$ tale che

$r_{\text{birth}} < r_{\text{death}}$, dove: r_{birth} è il raggio al quale nasce una classe omologica indipendente in $H_k(S(X, \sim_r))$; r_{death} è il raggio al quale quella classe diventa banale (confine di una catena in dimensione $k+1$). I punti con $r_{\text{death}} = \infty$ corrispondono a classi che non muoiono mai

nella filtrazione. La persistenza di un punto (b, d) è $d - b \geq 0$.

Il diagramma di persistenza cattura l'intera storia topologica della filtrazione. Ogni punto del diagramma corrisponde a una 'caratteristica topologica' che nasce e muore durante la filtrazione. Per PD_0 , le caratteristiche sono le componenti connesse: nascono tutte al raggio $r = 0$ (ogni elemento è da solo) e muoiono quando si fondono con altre componenti al crescere di r . Per PD_1 , le caratteristiche sono i cicli: nascono quando si forma un ciclo nel grafo e muoiono quando quel ciclo viene 'riempito' da triangoli.

La regola di nascita e morte per le componenti connesse in PD_0 è descritta dall'algoritmo del 'elder rule': quando due componenti si fondono (perché appare un arco tra di esse), la componente 'più giovane' (quella nata più tardi, ovvero quella con r_{birth} più grande) 'muore'. La componente 'più vecchia' sopravvive con il suo stesso r_{birth} . L'ultima componente sopravvissuta ha $r_{\text{death}} = \infty$ e il suo r_{birth} è il raggio minimo tale che il grafo è connesso (cioè il sorites threshold r^* con segno invertito: l'unica componente nata per ultima è quella che persiste all'infinito).

Esempio (predicato 'è alto'). Si consideri un insieme X di 5 stature campione: {163, 170, 175, 180, 187} cm con la metrica euclidea. Le distanze sono: $d(163,170)=7$, $d(170,175)=5$, $d(175,180)=5$, $d(180,187)=7$, $d(163,175)=12$, $d(170,180)=10$, $d(175,187)=12$, $d(163,180)=17$, $d(170,187)=17$, $d(163,187)=24$. La filtrazione di Sorite produce i seguenti eventi: a $r=5$, appaiono gli archi (170,175) e (175,180) (due componenti si fondono); a $r=7$, appaiono gli archi (163,170) e (180,187) (altre fusioni); a $r=10$, appare l'arco (170,180) (crea un triangolo, riempie il ciclo 170-175-180); ecc. Il diagramma PD_0 ha 5 punti, uno con $r_{\text{death}}=\infty$ e quattro con r_{death} finito. Il sorites threshold è $r^* = 24$ (il raggio minimo per cui la componente di 163 e quella di 187 si fondono).

9.3 Definizione formale del sorites barcode

Definizione 20 (Sorites threshold).

Il sorites threshold r^* è il valore del raggio al quale la componente connessa di $\min(X)$ e la componente connessa di $\max(X)$ si fondono per la prima volta nella filtrazione di Sorite. Formalmente:

$$r^* = \inf\{r \geq 0 : \min(X) \text{ e } \max(X) \text{ sono nella stessa componente connessa}\}$$

di $G(X, \sim_r)$ Per $r < r^*$, $\min(X)$ e $\max(X)$ sono in componenti diverse e non esiste

assegnamento coerente con contrasto (Corollario 1.1). Per $r \geq r^*$,
i

due estremi sono nella stessa componente e ogni assegnamento coerente è costante su di essa.

Definizione 21 (Sorites barcode).

Il sorites barcode del predicato P su $(X, d, \leq)$ è il diagramma di

persistenza PD_0 arricchito con: (i) L'indicazione, per ogni barra $(r_{\text{birth}}, r_{\text{death}})$, se la corrispondente componente contenga $\min(X)$ o $\max(X)$ (o nessuno dei due); (ii) Il sorites threshold r^* ;

(iii) Per le filtrazioni continue, anche i diagrammi PD_k per $k \geq 1$,

che descrivono la struttura ciclica del predicato. Il barcode arricchito è la firma topologica completa del predicato.

Il sorites barcode è la rappresentazione canonicamente stabile della struttura di vaghezza del predicato. Le barre con lunga persistenza corrispondono a caratteristiche topologiche 'robuste' — strutture di indistinguibilità che persistono per un ampio intervallo di soglie di tolleranza — mentre le barre con breve persistenza corrispondono a caratteristiche 'fragili' che dipendono dai dettagli specifici della metrica.

Nella pratica, il sorites barcode si visualizza come un insieme di barre orizzontali su un asse che rappresenta r : ogni barra si estende dal suo r_{birth} al suo r_{death} . Il sorites threshold r^* appare come il punto in cui la barra della componente di $\max(X)$ si fonde con la barra della componente di $\min(X)$, creando una barra infinita. Barre molto persistenti prima di r^* indicano la presenza di 'cluster' robusti di istanze simili — zone del dominio del predicato in cui gli elementi sono molto simili tra loro ma ben separati dagli altri cluster.

9.4 La relazione tra r^* e ε^*

Teorema 11 (Relazione tra sorites threshold e numero di risoluzione). Per uno spazio metrico finito (X, d) con l'ordinamento indotto dalla distanza da $\min(X)$ (cioè $x < y$ se $d(\min(X), x) < d(\min(X), y)$), il numero di risoluzione e il sorites threshold sono legati da:

$$\varepsilon^*(\sim_{r^*}) \cdot L_{\text{mono_min}}(\sim_{r^*}) = 1$$

dove $\sim_{r^*}$ è la relazione di tolleranza al raggio r^* .

Dimostrazione. Per definizione, r^* è il raggio minimo tale che $G(X, \sim_{r^*})$ è connesso con $\min(X)$ e $\max(X)$ nella stessa componente. Per il Teorema 3, $\varepsilon^*(\sim_{r^*}) =$

$1/L_mono_min(\sim_r^*)$). La relazione $\varepsilon^* \cdot L_mono_min = 1$ è dunque la stessa uguaglianza espressa diversamente. Il contenuto informativo del teorema è che il sorites threshold r^* è raggiunto esattamente quando $L_mono_min(\sim_r) \cdot r = r^* \cdot L_mono_min(\sim_r^*)$: ogni arco del cammino monotono minimo ha lunghezza esattamente r^* (altrimenti il cammino potrebbe essere trovato a un raggio minore, contraddicendo la minimalità di r^*). $\square$

Il Teorema 11 fornisce un'interpretazione fisica del sorites threshold: r^* è il raggio 'critico' di tolleranza, il punto in cui la struttura di indistinguibilità è appena sufficiente a collegare gli estremi del dominio del predicato. Per $r < r^*$, il paradosso di Sorite non si applica (gli estremi sono distinti in modo coerente); per $r \geq r^*$, il paradosso è completo (ogni classificazione coerente è costante).

Questa interpretazione ha un corrispettivo semantico preciso: r^* è la 'risoluzione critica' del predicato, la soglia di indistinguibilità al di sopra della quale il predicato diventa irrimediabilmente vago. Diversi parlanti della stessa lingua possono avere soglie di indistinguibilità diverse (riflettendo acuità percettive o norme linguistiche diverse), e il sorites threshold r^* aggregato a livello di comunità è la soglia minima che produce vaghezza collettiva.

9.5 Stabilità del sorites barcode

Teorema 12 (Stabilità del sorites barcode, Cohen-Steiner, Edelsbrunner, Harer 2007). Siano (X, d) e (X, d') due metriche sullo stesso insieme X con distanza uniforme $\|d - d'\|_\infty = \max_{\{x,y\}} |d(x,y) - d'(x,y)| \leq \delta$. La distanza bottleneck tra i diagrammi di persistenza soddisfa:

$$d_B(PD_k(d), PD_k(d')) \leq \delta \text{ per ogni } k \geq 0$$

dove la distanza bottleneck d_B tra due diagrammi di persistenza α e β è:

$$d_B(\alpha, \beta) = \inf_{\{\sigma : \alpha \rightarrow \beta \text{ bijection}\}} \sup_{\{p \in \alpha\}} \|p - \sigma(p)\|_\infty$$

dove la biiezione σ è tra le due copie del diagramma (con i punti della diagonale interpretati come punti di persistenza zero che possono essere accoppiati con punti reali o viceversa).

Dimostrazione. Cohen-Steiner, Edelsbrunner e Harer (2007), 'Stability of Persistence Diagrams', Discrete & Computational Geometry 37(1), 103-120. La dimostrazione usa la teoria dell'interleaving degli omorfismi di moduli persistenti: una perturbazione di δ nella metrica produce una perturbazione di δ in ciascun raggio di nascita/morte, che per la teoria dell'interleaving si traduce in una distanza bottleneck di al più δ . $\square$

Il Teorema 12 è il risultato di stabilità fondamentale per le applicazioni del sorites barcode. Garantisce che piccole perturbazioni della struttura di tolleranza del predicato (dovute, ad esempio, a errori di misura o a variazioni inter-individuali nelle soglie di indistinguibilità) producono piccole

perturbazioni nel barcode. Questa robustezza è essenziale per le applicazioni empiriche: se si estrae il barcode da dati sperimentali con errori di δ cm nella misura delle stature, il barcode vero e quello stimato differiscono di al più δ .

La distanza bottleneck d_B è la metrica naturale sullo spazio dei diagrammi di persistenza. È una metrica stabile (invariante rispetto a piccole perturbazioni) e computazionalmente trattabile: si calcola come il costo del matching ottimale tra i due diagrammi, che si riduce a un problema di trasporto ottimale su insiemi finiti di punti.

Corollario 12.1 (Stabilità del sorites threshold). Sotto le ipotesi del Teorema 12, il sorites threshold soddisfa $|r^*(d) - r^*(d')| \leq \delta$.

Dimostrazione. Il sorites threshold r^* è la persistenza massima della barra in PD_0 relativa alla componente di $\max(X)$. La distanza bottleneck $\leq \delta$ garantisce che questa persistenza non cambia di più di δ sotto la perturbazione della metrica. $\square$

9.6 Il diagramma PD_1 e la vaghezza ciclica persistente

Definizione 22 (Vaghezza ciclica persistente di ordine k).

La vaghezza ciclica persistente di ordine k del predicato P è misurata dal diagramma di persistenza PD_k della filtrazione di Sorite. Ogni

punto $(r_0, r_1) \in PD_k$ corrisponde a un ciclo k -dimensionale che:

nasce al raggio r_0 (il raggio al quale appare il primo arco/simpleso che crea il ciclo); muore al raggio r_1 (il raggio al quale compaiono abbastanza simplessi di dimensione $k+1$ da riempire il ciclo). La persistenza $r_1 - r_0$ misura la robustezza del paradosso ciclico corrispondente.

Per $k = 0$, la vaghezza ciclica persistente è la vaghezza standard misurata da r^* : il ciclo di ordine 0 è la componente connessa che nasce e muore mentre le componenti di $\min(X)$ e $\max(X)$ si fondono.

Per $k = 1$, ogni punto (r_0, r_1) in PD_1 corrisponde a un ciclo del grafo di tolleranza che persiste nell'intervallo $[r_0, r_1)$. Un ciclo con grande persistenza $r_1 - r_0$ indica una struttura di indistinguibilità ciclica molto robusta: il ciclo resiste all'aumento della soglia di tolleranza per un ampio intervallo, prima di essere 'riempito' da triangoli. Questo corrisponde a una zona del dominio del predicato in cui il paradosso ciclico è particolarmente pronunciato e difficile da eliminare.

Interpretazione della vaghezza di second'ordine. I cicli in PD_1 misurano la 'vaghezza sulla vaghezza': non soltanto è incerto dove passa il confine tra Vero e Falso, ma è incerto dove si trovano le zone di incertezza stessa. Un predicato con molte barre in PD_1 ha una struttura di vaghezza ramificata: ci sono molte 'isole di ambiguità' circoscritte, ciascuna delle quali produce il suo paradosso ciclico locale.

Esempio. Per il predicato 'è di sinistra' su un dominio con struttura circolare (spettro politico chiuso), il diagramma PD_1 ha una barra di persistenza molto lunga: il ciclo fondamentale del cerchio nasce al raggio di tolleranza minimo (quando le posizioni adiacenti divengono indistinguibili) e non muore mai (perché il cerchio non ha un 'interno' 2-dimensionale che potrebbe riempire il ciclo). Questo corrisponde alla vaghezza ciclica irrisolvibile del predicato descritta nel Capitolo 6.

Per il predicato 'è bello' su un dominio bidimensionale (con dimensioni di 'attrattività facciale' e 'attrattività fisica'), il diagramma PD_1 ha barre di persistenza intermedie: i cicli nascono quando si formano anelli di indistinguibilità nella griglia bidimensionale e muoiono quando la soglia di tolleranza è sufficientemente alta da riempire i quadrati elementari con triangoli. La struttura di PD_1 rivela la 'topologia della vaghezza' del predicato in senso letterale: la forma degli 'spazi vuoti' nella rete di indistinguibilità.

9.7 Il sorites barcode per predicati multidimensionali

Per i predicati genuinamente multidimensionali, il sorites barcode ha una struttura molto più ricca di quella dei predicati unidimensionali. In particolare, il diagramma PD_1 può avere molte barre con persistenza variabile, e la struttura del barcode è un invariante dello spazio dei dati che non può essere ridotta a informazione unidimensionale.

Proposizione 20 (Sorites barcode della griglia). Per la griglia $m \times n$ con distanza euclidea, il diagramma PD_1 ha $(m-1)(n-1)$ punti, tutti con $r_{\text{birth}} = 1$ e $r_{\text{death}} = \sqrt{2}$ (per $m, n \geq 2$). Tutti i cicli nascono allo stesso raggio (quando si formano i quadrati elementari del grafo) e muoiono allo stesso raggio (quando appare la diagonale del quadrato, creando un triangolo che riempie il ciclo).

Dimostrazione. I cicli della griglia sono tutti i cicli quadrati elementari. Un ciclo quadrato $(i,j)-(i+1,j)-(i+1,j+1)-(i,j+1)$ nasce al raggio $r=1$ (quando appaiono tutti e 4 gli archi di lunghezza 1 del quadrato; l'arco 'critico' che completa il ciclo è l'ultimo dei quattro) e muore al raggio $r=\sqrt{2}$ (quando appare la diagonale del quadrato, di lunghezza $\sqrt{2}$, che crea il triangolo $\{(i,j),(i+1,j),(i+1,j+1)\}$ o $\{(i,j),(i,j+1),(i+1,j+1)\}$ che riempie il ciclo). La coincidenza di r_{birth} e r_{death} per tutti i cicli produce un diagramma PD_1 'monocromatico': tutte le barre hanno la stessa lunghezza. $\square$

Il fatto che tutti i cicli abbiano la stessa persistenza $\sqrt{2} - 1 \approx 0.414$ per la griglia con distanza euclidea riflette la simmetria della struttura della griglia. Per predicati multidimensionali con

struttura di tolleranza non uniforme (alcune dimensioni più 'variabili' di altre), le barre di PD_1 avranno persistenze diverse, e la distribuzione delle persistenze nel diagramma rivela la struttura di anisotropia del predicato.

Proposizione 21 (Confronto di predicati tramite il barcode). Dati due predicati P e P' con sorites barcode $B(P)$ e $B(P')$, la loro distanza è $d_B(B(P), B(P')) \leq |d_P - d_{P'}|_\infty$, dove d_P e $d_{P'}$ sono le metriche sui rispettivi domini (normalizzate al diametro 1). La distanza bottleneck fornisce quindi un indice quantitativo di 'diversità di struttura di vaghezza' tra due predicati.

Dimostrazione. Diretta dal Teorema 12 (stabilità), applicato alla differenza tra le due metriche. $\square$

Questa proposizione ha applicazioni immediate nella linguistica computazionale: permette di confrontare quantitativamente la struttura di vaghezza di predicati in lingue diverse, o in varietà diverse della stessa lingua. Se due lingue hanno barcodes molto simili per un certo predicato (ad esempio 'è alto'), questo suggerisce che la struttura di indistinguibilità è culturalmente o biologicamente determinata (struttura universale); se i barcodes differiscono significativamente, la struttura di vaghezza è culturalmente specifica.

Capitolo 10. L'Algebra degli Assegnamenti Ottimali

I capitoli precedenti hanno analizzato gli assegnamenti di valori di verità dal punto di vista dell'omologia (che caratterizza le ostruzioni globali) e del numero di risoluzione (che misura il costo minimo di quasi-coerenza). Questo capitolo sviluppa una prospettiva algebrica sugli assegnamenti, usando la coomologia simpliciale del complesso di Sorite per caratterizzare la struttura dello spazio degli assegnamenti ε -coerenti. Il risultato principale è che lo spazio degli assegnamenti ottimali ha la struttura di un poliedro convesso, e la sua geometria è direttamente collegata alla struttura topologica del complesso di Sorite.

10.1 L'algebra di tolleranza

Definizione 24 (Algebra di tolleranza).

Dato uno spazio di tolleranza finito $(X, \sim)$, l'algebra di tolleranza $A(X, \sim)$ è lo spazio vettoriale $\mathbb{R}^X$ dotato del prodotto: $(f \otimes g)(x) = g(x) \cdot (T_{\sim} f)(x) = g(x) \cdot \sum_{y \sim x} f(y) / \deg(x)$. Questo prodotto rappresenta la 'convoluzione di f lungo i vicini di tolleranza, pesata da g al punto corrente'. Non è commutativo in generale.

L'algebra di tolleranza codifica in modo compatto la struttura delle interazioni tra assegnamenti: il prodotto $f \circledast g$ misura come il valore di g in un punto x interagisce con la media dei valori di f nei vicini di x . Questo prodotto emerge naturalmente nello studio dei processi di aggiornamento delle credenze in un sistema di agenti razionali con struttura di tolleranza.

Proposizione 22 (Struttura modulare dell'algebra di tolleranza). L'algebra $A(X, \sim)$ è un modulo a sinistra su $\mathbb{R}$, compatibile con la struttura spettrale di $T_{\sim}$. Gli autospazi di $T_{\sim}$ sono sottoalgebre di $A(X, \sim)$: se f è un autovettore di $T_{\sim}$ con autovalore λ , allora $f \circledast g = \lambda \cdot fg$ per ogni g (dove fg è il prodotto puntuale). In particolare, gli assegnamenti coerenti (autovettori di autovalore 1) formano una sottoalgebra rispetto al prodotto puntuale.

Dimostrazione. Se $T_{\sim} f = \lambda f$, allora $(f \circledast g)(x) = g(x) \cdot (T_{\sim} f)(x) = g(x) \cdot \lambda f(x) = \lambda(fg)(x)$. La sottoalgebra degli assegnamenti coerenti ($\lambda=1$) è chiusa rispetto al prodotto puntuale perché il prodotto di due assegnamenti coerenti è ancora coerente (la coerenza è una condizione su archi, e il prodotto puntuale di due funzioni costanti su ogni componente è ancora costante). $\square$

10.2 Coomologia degli assegnamenti

La coomologia simpliciale del complesso di Sorite fornisce una prospettiva duale all'omologia: invece di catene (combinazioni lineari di semplici) e bordi, si lavora con cocatene (funzioni sui semplici) e cobordi. Gli assegnamenti coerenti hanno una naturale interpretazione come 0-cocicli.

Definizione 25 (Spazio delle cocatene).

Una k -cocatena è una funzione φ che assegna un valore reale a ogni

k -simpleso di $S(X, \sim)$. Lo spazio delle k -cocatene è $C^k(S(X, \sim); \mathbb{R}) \cong \mathbb{R}^{f_k}$, dove f_k è il numero di k -simplessi del complesso.

Definizione 26 (Cobordo).

Il cobordo $\delta^k: C^k \rightarrow C^{k+1}$ è definito per una k -cocatena φ e un

$(k+1)$ -simpleso $\sigma = \{x_0, \dots, x_{k+1}\}$ da:

$$(\delta^k \varphi)(\sigma) = \sum_{i=0}^{k+1} (-1)^i \cdot \varphi(\{x_0, \dots, x_i, \dots, x_{k+1}\})$$

dove il cappello indica l'omissione dell'elemento. Il cobordo è il

duale del bordo ∂ .

Proposizione 23 (Assegnamenti coerenti come 0-cocicli). Un assegnamento $f: X \rightarrow \mathbb{R}$ è coerente se e solo se $\delta^0 f = 0$, dove $\delta^0 f$ è la 1-cocatena definita da $(\delta^0 f)(x, y) = f(y) - f(x)$ per ogni arco (x, y) orientato del grafo di tolleranza.

Dimostrazione. La condizione di coerenza $f(x) = f(y)$ per ogni coppia tollerante $x \sim y$ equivale a $f(y) - f(x) = 0$ per ogni arco (x, y) , che è esattamente la condizione $(\delta^0 f)(x, y) = 0$ per ogni arco. $\square$

Proposizione 24 (Caratterizzazione coomologica degli assegnamenti non costanti). Esiste un assegnamento coerente non costante $f: X \rightarrow \mathbb{R}$ se e solo se $H^0(S(X, \sim); \mathbb{R})$ ha rango ≥ 2 , ovvero se e solo se $G(X, \sim)$ è disconnesso. Questo rafforza il Corollario 1.1: l'impossibilità di un assegnamento coerente con contrasto è esattamente l'impossibilità di avere rango ≥ 2 in H^0 mantenendo la connessione.

Dimostrazione. $H^0(S(X, \sim); \mathbb{R}) = \ker(\delta^0) / \text{im}(\partial_1^*) = \ker(\delta^0)$ (poiché non ci sono (-1) -simplessi, non c'è nulla nell'immagine). Il nucleo di δ^0 è lo spazio degli assegnamenti coerenti. La dimensione di questo nucleo è uguale al numero di componenti connesse del grafo. Dunque un assegnamento coerente non costante esiste se e solo se il numero di componenti è ≥ 2 . $\square$

10.3 La struttura degli assegnamenti ε -coerenti come poliedro

Proposizione 25 (Poliedro degli assegnamenti ε -coerenti). L'insieme P_ε degli assegnamenti ε -coerenti monotoni con $f(\min) = 0$ e $f(\max) = 1$ è un poliedro convesso compatto in $\mathbb{R}^n$, definito da:

$$P_\varepsilon = \{f \in [0, 1]^n : f(x_i) \leq f(x_j) \text{ per } i \leq j, |f(x_i) - f(x_j)| \leq \varepsilon \text{ per } x_i \sim x_j, f(x_1) = 0, f(x_n) = 1\}$$

Dimostrazione. Ogni vincolo è un'uguaglianza o disuguaglianza lineare in f , dunque P_ε è un poliedro. È compatto perché è contenuto in $[0, 1]^n$. Per il Teorema 3, P_ε è non vuoto se e solo se $\varepsilon \geq \varepsilon^* = 1/L_{\text{mono_min}}$. $\square$

Proposizione 26 (Geometria del poliedro P_ε). Per $\varepsilon > \varepsilon^*$, il poliedro P_ε è non vuoto e ha dimensione positiva (contiene un aperto relativo). Per $\varepsilon = \varepsilon^*$, $P_{\{\varepsilon^*\}}$ è non vuoto ma ha dimensione inferiore: è un poliedro 'piatto' (meno gradi di libertà). I vertici di P_ε (i punti estremi del poliedro) sono assegnamenti con la maggiore specificità possibile: la funzione $f(y_j) = j/L$ sui nodi del cammino e valori 'estremi' sugli altri nodi.

Dimostrazione. La non-vacuità e la dimensione positiva per $\varepsilon > \varepsilon^*$ seguono dal fatto che la costruzione della Parte 1 del Teorema 3 produce un punto interno all'interno di P_ε : ogni piccola perturbazione di f rimane in P_ε per ε sufficientemente più grande di ε^* . La struttura dei vertici segue dall'analisi degli assegnamenti che saturano i vincoli di ε -coerenza (cioè per cui $|f(x_i) - f(x_j)| = \varepsilon$ per ogni coppia tollerante critica). $\square$

Corollario 10.1 (Il 'costo di ambiguità'). Il volume di P_ε cresce monotonamente con $\varepsilon - \varepsilon^*$. La crescita del volume misura quante 'scelte' di funzione di membership sono disponibili al di sopra del costo minimo ε^* : un volume grande indica molta ambiguità residua anche dopo aver fissato il costo di coerenza, mentre un volume piccolo indica che la struttura di tolleranza determina quasi univocamente la funzione di membership anche per ε solo leggermente superiore a ε^* .

10.4 Assegnamenti ottimali e funzioni di membership nelle logiche fuzzy

Proposizione 27 (Giustificazione della funzione di membership lineare, Łukasiewicz). Per una tolleranza di larghezza uniforme k su $X = \{1, \dots, n\}$, la funzione di membership ottimale — quella che minimizza ε sulla funzione ε -coerente monotona con condizioni agli estremi — è unica (sui nodi del cammino monotono minimo) e uguale a $f(x_i) = (i-1)/(n-1)$. Questa è la funzione di membership lineare della logica fuzzy di Łukasiewicz.

Dimostrazione. Il cammino monotono minimo per la tolleranza di larghezza uniforme k è $x_1, x_{1+k}, x_{1+2k}, \dots, x_n$ (la sequenza che salta k posizioni per volta). Per la Proposizione 8 del Capitolo 5, i valori sui nodi del cammino sono univocamente determinati: $f(x_{1+jk}) = j/L_{\text{mono_min}} = j \cdot k / (n-1)$. Poiché $j \cdot k / (n-1) = (1+jk-1)/(n-1)$ è la funzione lineare $(i-1)/(n-1)$ valutata in $i = 1+jk$, i valori dell'assegnamento ottimale sui nodi del cammino coincidono con la funzione lineare. Per gli elementi fuori dal cammino (quelli tra due nodi consecutivi del cammino), la funzione di membership è vincolata dai valori adiacenti e dalla monotonia a seguire la funzione lineare (per la struttura uniforme della tolleranza, la funzione lineare è l'unico modo di interpolare monotonamente e ε^* -coerentemente tra i nodi del cammino). $\square$

Questo risultato è notevole: la funzione di membership lineare di Łukasiewicz, spesso criticata come arbitraria e priva di motivazione intrinseca, è in realtà l'unica funzione che ottimizza simultaneamente monotonia e quasi-coerenza per predicati con struttura di tolleranza uniforme. La 'linearità' non è una scelta arbitraria ma una conseguenza necessaria della struttura combinatoria del predicato. Per predicati con struttura di tolleranza non uniforme, la funzione di membership ottimale sarà non lineare, e la sua forma riflette la struttura specifica delle relazioni di indistinguibilità.

Il quadro dell'algebra degli assegnamenti ottimali fornisce quindi una risposta parziale al problema della 'scelta della funzione di membership' nelle logiche fuzzy: la funzione ottimale è determinata (almeno sui punti del cammino critico) dalla struttura di tolleranza del predicato, e non richiede

scelte arbitrarie. La 'parte libera' della funzione di membership — i valori negli elementi fuori dal cammino critico — rimane indeterminata, e questa indeterminazione corrisponde alla genuina ambiguità semantica residua del predicato.

Capitolo 11. Equivalenza tra Predicati e Classificazione

Una teoria matematica della vaghezza deve fornire strumenti per confrontare e classificare predicati diversi. Due predicati 'sono ugualmente vaghi'? Quando un predicato è 'più vago' di un altro? Queste domande richiedono nozioni precise di equivalenza e distanza tra predicati. Questo capitolo sviluppa la teoria delle equivalenze tra predicati (basata sugli isomorfismi di spazi di tolleranza) e degli invarianti completi (basata sul sorites barcode).

11.1 Morfismi e equivalenze di predicati

Definizione 27 (Equivalenza sorite).

Due predicati P e P' su spazi di tolleranza $(X, \sim)$ e $(X', \sim')$ sono

sorites-equivalenti se esiste una biiezione $\varphi: X \rightarrow X'$ che preserva la

relazione di tolleranza biunivocamente: $x \sim y$ se e solo se $\varphi(x) \sim'$

$\varphi(y)$.

L'equivalenza sorite è l'equivalenza 'minimale' che preserva tutta la struttura di vaghezza: due predicati sorites-equivalenti hanno esattamente la stessa struttura combinatoria di indistinguibilità. In particolare, hanno lo stesso numero di risoluzione ε^* , lo stesso sorites barcode, e le stesse ostruzioni omologiche.

Teorema 13 (Invarianza del sorites barcode sotto equivalenza sorite). Se $(X, \sim)$ e $(X', \sim')$ sono sorites-equivalenti tramite φ , i loro sorites barcode sono isometricamente equivalenti nel senso della distanza bottleneck: $d_B(B(X, \sim), B(X', \sim')) = 0$.

Dimostrazione. φ induce un isomorfismo di grafi $G(X, \sim) \rightarrow G(X', \sim')$, un isomorfismo di complessi simpliciali $S(X, \sim) \rightarrow S(X', \sim')$, e dunque isomorfismi $H_k(S(X, \sim)) \cong H_k(S(X', \sim'))$ per ogni k . La filtrazione di Sorite di (X', d') dove $d'(\varphi(x), \varphi(y)) = d(x, y)$ coincide con quella di (X, d) : le due filtrazioni sono identiche a meno del rinominamento dei vertici. Dunque i diagrammi di persistenza sono identici. $\square$

11.2 Il sorites barcode come invariante completo: risultati e limiti

Teorema 14 (Classificazione completa per alberi). Il sorites barcode PD_0 classifica completamente i predicati finiti il cui grafo di tolleranza è un albero, a meno di equivalenza sorite.

Dimostrazione. Un albero finito pesato è completamente determinato dalla sua struttura ultrametrica: la distanza ultrametrica tra due vertici è il massimo peso degli archi nel cammino unico che li connette. Il diagramma PD_0 della filtrazione di Vietoris-Rips di un albero codifica esattamente questa struttura ultrametrica: ogni barra $(0, w)$ in PD_0 corrisponde a un arco di peso w nell'albero (l'arco che, quando compare nella filtrazione, fonde due componenti). Carlsson e Mémoli (2010) dimostrano che questa corrispondenza è biunivoca: PD_0 determina l'albero a meno di isomorfismo, che è la stessa cosa dell'equivalenza sorite. $\square$

Proposizione 28 (Non completezza per grafi con cicli). Il sorites barcode completo (tutti i PD_k) non classifica in generale i predicati il cui grafo di tolleranza contiene cicli.

Dimostrazione. Si costruisce un controesempio con $|X| = 5$. Siano G_1 il ciclo C_5 con un vertice aggiuntivo connesso a uno dei 5 vertici del ciclo, e G_2 il grafo formato da due triangoli (K_3) che condividono un arco. Entrambi hanno $|V| = 6$, $|E| = 6$, e 1 componente connessa, dunque $rk(H_1) = |E| - |V| + 1 = 1$ per entrambi. Il diagramma PD_0 è identico (un albero di fusione delle 6 componenti) e PD_1 ha una barra di uguale persistenza (per un'opportuna scelta dei pesi degli archi). Tuttavia G_1 e G_2 non sono isomorfi (G_1 ha massimo grado 3, G_2 ha massimo grado 4), dunque non sono sorites-equivalenti. $\square$

11.3 Invarianti più fini: profilo di Betti e persistenza filtrata

Definizione 28 (Profilo di Betti della filtrazione).

Il profilo di Betti di ordine k è la funzione $\beta_k(r) = rk(H_k(S(X,$

$\sim_r)))$ che misura il numero di 'buchi' k -dimensionali nel complesso di Sorite al raggio r . Il profilo di Betti è meno informativo del diagramma di persistenza (non cattura il raggio esatto di nascita e morte), ma è più semplice da calcolare e da interpretare graficamente.

Proposizione 29 (Ricostruzione del profilo di Betti dal diagramma di persistenza). Il profilo di Betti si ricostruisce dal diagramma di persistenza come:

$$\beta_k(r) = |\{(b,d) \in PD_k : b \leq r < d\}|$$

Dimostrazione. Diretta dalla definizione del diagramma di persistenza come insieme dei punti di nascita e morte delle classi omologiche. $\square$

Il profilo di Betti $\beta_0(r)$ è particolarmente intuitivo: misura il numero di 'cluster' distinti di istanze simili al livello di tolleranza r . Per r piccolo, ogni istanza forma il suo cluster ($\beta_0 = n$); per $r \geq r^*$, tutte le istanze sono in un unico cluster ($\beta_0 = 1$). Il profilo di Betti $\beta_0(r)$ descrive quindi la struttura gerarchica di clustering del predicato: come le istanze si raggruppano progressivamente al crescere della soglia di tolleranza.

Il profilo di Betti $\beta_1(r)$ misura il numero di 'cicli' nel grafo di tolleranza al raggio r . Un picco in $\beta_1(r)$ indica un intervallo di soglie in cui si formano molti cicli (zone del dominio con alta densità di relazioni di tolleranza che producono strutture cicliche) senza che questi cicli vengano riempiti da triangoli. Il profilo $\beta_1(r)$ rivela la struttura 'anulare' del predicato: le zone in cui la vaghezza ha una natura ciclica piuttosto che lineare.

11.4 Distanza tra predicati e spazio delle strutture di vaghezza

Definizione 29 (Spazio dei predicati con metrica di sorites).

Lo spazio $\mathcal{F}$ di tutti i predicati finiti (spazi di tolleranza finiti con ordinamento e metrica) è dotato della pseudometrica:

$$d(P, P') = d_B(B(P), B(P'))$$

dove $B(P)$ è il sorites barcode completo di P . Due predicati sono alla distanza 0 se e solo se hanno lo stesso barcode.

Proposizione 30 (Proprietà metriche di $d(P, P')$). d definisce una pseudometrica sullo spazio $\mathcal{F}$: è simmetrica, soddisfa $d(P, P) = 0$, e verifica la disuguaglianza triangolare. Non è una metrica vera perché $d(P, P') = 0$ non implica P sorites-equivalente a P' (per i grafi con cicli, due predicati non isomorfi possono avere lo stesso barcode).

Dimostrazione. Le proprietà metriche di d_B si ereditano direttamente. La non-injectività è dimostrata dalla Proposizione 28. $\square$

La pseudometrica d fornisce un modo quantitativo per rispondere alla domanda 'quanto sono vaghi due predicati?': due predicati molto vicini nella metrica d hanno strutture di vaghezza quasi identiche, e la distanza misura esattamente quanto devono essere perturbate le loro metriche di indistinguibilità per render i due predicati strutturalmente identici.

Applicazione. Si confrontino le strutture di vaghezza del predicato 'è alto' in italiano e in inglese ('is tall'), estratte da corpora paralleli. Se i barcodes differiscono significativamente, questo suggerisce che la struttura di indistinguibilità semantica per l'altezza è diversa nelle due comunità linguistiche: le soglie di JND percettiva sono uguali (fisiologicamente determinate), ma le norme linguistiche producono strutture di aggregazione diverse. Una distanza $d(P_{\text{it}}, P_{\text{en}}) > 0$ misura la 'diversità cross-linguistica della vaghezza' del predicato, mentre $d = 0$ (o molto piccolo) indicherebbe universalità semantica.

Capitolo 12. Vaghezza e Logiche Formali: Connessioni Strutturali

Le risposte filosofiche al paradosso di Sorite — epistemicismo, supervalutazione, logiche a gradi — corrispondono, nel quadro matematico sviluppato in questo saggio, a scelte precise sulla struttura topologica del dominio del predicato. Questo capitolo rende esplicite queste corrispondenze, mostrando come ciascuna tradizione filosofica possa essere 'tradotta' nel linguaggio degli spazi di tolleranza e dei sorites barcode. La matematica non risolve il dibattito filosofico, ma lo chiarisce: mostra esattamente quali ipotesi strutturali corrispondono a ciascuna posizione filosofica, e misura quantitativamente il 'costo' di ciascuna posizione.

12.1 L'epistemicismo e il numero di risoluzione

L'epistemicismo sostiene che ogni predicato vago ha un confine netto: esiste un intero n^* tale che n^* grani non formano un mucchio ma n^*+1 grani sì. La nostra incapacità di determinare n^* è una limitazione epistemica, non una caratteristica del predicato. Nel linguaggio degli spazi di tolleranza, questa posizione equivale all'affermazione che il grafo di tolleranza 'vero' del predicato — quello determinato dall'uso completo della comunità linguistica nel tempo — è disconnesso. La relazione di tolleranza osservata empiricamente è un'approssimazione rumorosa della relazione vera.

Il costo dell'epistemicismo è misurato dalla distanza bottleneck tra il barcode del predicato osservato e quello del predicato 'epistemico' (con confine netto). Un predicato epistemico ha PD_0

con soltanto due componenti (la componente di $\min(X)$ e quella di $\max(X)$) e $r^* = 0$. La distanza bottleneck misura quanto lontano è il predicato osservato da questa struttura ideale.

Proposizione 31 (Costo epistemico). Il costo epistemico di un predicato P — la distanza bottleneck tra $B(P)$ e il barcode di un predicato epistemico (con confine netto) — è uguale al sorites threshold $r^*(P)$: il raggio di tolleranza minimo che connette $\min(X)$ e $\max(X)$. Predicati con r^* grande sono 'lontani' dall'epistemicismo.

Dimostrazione. Un predicato epistemico ha $r^* = 0$: gli estremi sono in componenti diverse dalla filtrazione stessa (al raggio 0). Il barcode di tale predicato ha la barra di $\min(X)$ che non muore mai e la barra di $\max(X)$ che non muore mai. La distanza bottleneck tra questo barcode e quello di P è determinata dal raggio $r^*(P)$ al quale la barra di $\max(X)$ muore (si fonde con la barra di $\min(X)$): questa è la perturbazione minimale della metrica necessaria per rendere P epistemico. $\square$

L'interpretazione è la seguente: per rendere il predicato 'è alto' epistemico (cioè per trovare un confine netto), si dovrebbero 'accorciare' tutte le distanze tra istanze adiacenti nella zona di confine di almeno r^* cm (il sorites threshold del predicato). La quantità r^* misura quanto lontana è la struttura di tolleranza osservata da quella di un predicato con confine netto.

12.2 La supervalutazione e lo spazio delle precisazioni

La teoria della supervalutazione (Fine, 1975) lavora con famiglie di precisazioni: ognuna fissa un confine netto, ma nessuna precisazione specifica è privilegiata. Nel linguaggio degli spazi di tolleranza, una precisazione corrisponde a un 'taglio' del grafo di tolleranza: la rimozione di uno o più archi che rende il grafo disconnesso, separando $\min(X)$ da $\max(X)$. La famiglia di tutte le precisazioni ammissibili corrisponde all'insieme di tutti i tagli minimali che separano $\min(X)$ da $\max(X)$.

Proposizione 32 (Struttura dello spazio delle precisazioni). Lo spazio delle precisazioni ammissibili — i tagli minimali di $G(X, \sim)$ che separano $\min(X)$ da $\max(X)$ — è in biiezione con i cammini monotoni minimi da $\min(X)$ a $\max(X)$ nel DAG $G^+(X, \sim)$. Ogni cammino monotono minimo corrisponde a una precisazione 'ottimale': la rimozione di tutti gli archi del cammino disconnette il grafo separando gli estremi.

Dimostrazione. Una precisazione minimale deve rimuovere un insieme di archi che (i) disconnette $G(X, \sim)$ separando $\min(X)$ da $\max(X)$, e (ii) è minimale con questa proprietà. Il taglio minimo (nel senso di max-flow min-cut) che separa $\min(X)$ da $\max(X)$ corrisponde agli archi del cammino critico. Ogni cammino monotono minimo attraversa un insieme di archi che, se rimossi, disconnette il grafo: la rimozione di tutti gli archi di un cammino monotono minimo è una precisazione minimale. $\square$

La struttura dello spazio delle precisazioni spiega un aspetto importante della critica alla supervalutazione. Fine (1975) afferma che l'enunciato 'esiste n tale che n grani non formano un mucchio ma $n+1$ grani sì' è supervero (vero in ogni precisazione), anche se nessuna istanza specifica è supervera. Nel quadro matematico, questo corrisponde al fatto che ogni cammino monotono minimo ha un arco critico (l'arco che attraversa il 'confine' della precisazione), ma archi critici diversi corrispondono a precisazioni diverse, e nessun arco specifico è 'l'arco critico di tutte le precisazioni'.

Il sorites barcode fornisce informazioni sulla distribuzione delle precisazioni nello spazio: le barre in PD_0 con lunga persistenza corrispondono a tagli 'robusti' (precisazioni stabili che rimangono valide per un ampio intervallo di perturbazioni), mentre le barre con breve persistenza corrispondono a tagli 'fragili' (precisazioni che dipendono da dettagli specifici della metrica e che cambierebbero con piccole perturbazioni).

12.3 Le logiche fuzzy e l'assegnamento ottimale

Le logiche a gradi di verità assegnano valori di membership continui in $[0,1]$. Il numero di risoluzione ε^* e la struttura degli assegnamenti ottimali (Proposizione 27) forniscono una risposta matematica alla domanda 'quale funzione di membership scegliere?'.

Il quadro matematico suggerisce la seguente gerarchia: (1) La funzione di membership 'minimalmente ε -incoerente' — quella che minimizza ε tra tutte le funzioni monotone con le condizioni agli estremi — è unica sui nodi del cammino monotono minimo (Proposizione 8), e per tolleranze uniformi coincide con la funzione lineare di Łukasiewicz (Proposizione 27). (2) Fuori dal cammino minimo, la funzione di membership ha gradi di libertà: lo spazio $P_{\{\varepsilon^*\}}$ è un poliedro che può avere dimensione positiva (Proposizione 26). (3) La funzione di Łukasiewicz (lineare) è un vertice 'canonico' del poliedro $P_{\{\varepsilon^*\}}$ per tolleranze uniformi.

Proposizione 33 (Ottimalità della funzione di Łukasiewicz per tolleranze uniformi). Per la tolleranza di larghezza uniforme k su $X = \{1, \dots, n\}$, la funzione $f(x_i) = (i-1)/(n-1)$ è l'unica funzione di membership che sia simultaneamente: (a) monotona non decrescente; (b) ε^* -coerente con $\varepsilon^* = k/(n-1)$; (c) simmetrica rispetto all'ordinamento inverso ($f(x_i) = 1 - f(x_{n+1-i})$). Le proprietà (a)-(c) determinano f univocamente.

Dimostrazione. (a) e (b) determinano f su tutti gli elementi: la monotonia e la ε^* -coerenza fissano $f(x_i) = (i-1)/(n-1)$ sui nodi del cammino (Proposizione 8), e per gli elementi fuori dal cammino la monotonia e la ε^* -coerenza combinati con la simmetria (c) lasciano soltanto la funzione lineare come possibilità. La simmetria rispetto all'ordinamento inverso cattura l'intuizione che 'calvo' e 'non-calvo' abbiano strutture di vaghezza simmetriche. $\square$

La proprietà (c) — la simmetria — è una condizione filosoficamente motivata: un predicato vago e la sua negazione dovrebbero avere strutture di vaghezza simmetriche. La logica fuzzy di

Łukasiewicz soddisfa questa proprietà con la negazione $f \mapsto 1-f$. Il fatto che la simmetria, combinata con la coerenza ottimale, determini la funzione lineare fornisce una giustificazione intrinseca alla scelta di Łukasiewicz.

12.4 Vaghezza higher-order e logiche di ordine superiore

La vaghezza higher-order è la vaghezza della vaghezza stessa: 'è x chiaramente alto?' può essere una domanda vaga. Nel linguaggio degli spazi di tolleranza, la vaghezza higher-order emerge dalla struttura del sorites barcode in gradi superiori: il diagramma PD_1 misura la vaghezza ciclica (ostruzioni a ordine 1), mentre un ipotetico PD_2 misurerebbe la vaghezza ciclica di secondo ordine.

La Congettura 4 del Capitolo 14 afferma che per predicati linguistici naturali le ostruzioni omologiche significative vivono soltanto in H_0 e H_1 (vaghezza di primo e secondo ordine). Questo corrisponde all'osservazione empirica che la vaghezza higher-order nei linguaggi naturali è di ordine finito: 'è x chiaramente chiaramente alto?' ha già una struttura di vaghezza che i parlanti tendono a trattare come equivalente a 'è x chiaramente alto?' per $k \geq 2$.

La struttura matematica suggerisce una gerarchia di logiche a gradi corrispondente alla gerarchia di ordini di vaghezza: la logica fuzzy di Łukasiewicz è adeguata per la vaghezza di primo ordine (ostruzioni in H_0), ma predicati con ostruzioni in H_1 richiederebbero logiche a valori in strutture topologiche più ricche (come il cerchio S^1 invece dell'intervallo $[0,1]$). Questa è una direzione di ricerca futura promettente, che connette la teoria degli spazi di tolleranza con la teoria delle logiche a valori in strutture algebriche non lineari.

Capitolo 13. Aspetti Computazionali

Questo capitolo descrive gli algoritmi per il calcolo degli invarianti principali introdotti in questo saggio — il numero di risoluzione ε^* , il sorites barcode, il vettore di Fiedler — e discute le tecniche sperimentali e computazionali per estrarre la struttura di tolleranza di un predicato da dati empirici. L'obiettivo è rendere la teoria non soltanto matematicamente rigorosa ma anche computazionalmente praticabile.

13.1 Algoritmi per il sorites barcode

Il calcolo del sorites barcode richiede tre passaggi principali: (1) la costruzione della filtrazione di Sorite (equivalentemente, della filtrazione di Vietoris-Rips); (2) il calcolo dell'omologia persistente per ogni complesso della filtrazione; (3) la lettura del diagramma di persistenza dalla riduzione della matrice dei bordi.

Librerie software efficienti per questo calcolo sono disponibili in diversi linguaggi di programmazione. Per Python: Gudhi (Geometric Understanding in Higher Dimensions), Ripser, e Giotto-TDA forniscono implementazioni ottimizzate della filtrazione di Vietoris-Rips e del calcolo dell'omologia persistente. Per Julia: Eirene.jl è particolarmente efficiente per grandi insiemi di dati. Per Java: JavaPlex offre funzionalità complete di TDA.

Proposizione 34 (Complessità computazionale del sorites barcode). Per un insieme X di n istanze del predicato: (a) La costruzione della filtrazione di Vietoris-Rips fino alla dimensione d richiede tempo $O(n^{d+1})$ nel caso peggiore. (b) Il calcolo dell'omologia persistente richiede la riduzione della matrice dei bordi, con complessità $O(m^3)$ nel caso peggiore dove m è il numero totale di semplici, ma $O(m^{\omega})$ in pratica (dove $\omega < 3$ è il coefficiente di complessità della moltiplicazione matriciale, e le strutture sparse riducono ulteriormente il costo). (c) Per $n \leq 1000$ e $d = 1$ (solo 0-simplici e 1-simplici), il calcolo è completato in pochi secondi su hardware moderno.

Dimostrazione. (a) Il numero di k -simplici in $VR(X, r^*)$ è al più $\beta(n, k+1)$, il coefficiente binomiale. Per $d = 1$, il numero di archi è $O(n^2)$. (b) La riduzione della matrice dei bordi è un'operazione di algebra lineare su $\mathbb{Z}/2\mathbb{Z}$; i risultati di Edelsbrunner, Letscher e Zomorodian (2002) e le implementazioni ottimizzate di Bauer (2021, Ripser) riducono il costo effettivo a $O(m \log m)$ per molti casi pratici. (c) Verifica sperimentale su dataset tipici per la linguistica. $\square$

13.2 Algoritmo per il numero di risoluzione ε^*

Il calcolo di ε^* è molto più semplice di quello del sorites barcode, e può essere eseguito in tempo lineare nel numero di archi del grafo di tolleranza.

Algoritmo per ε^* : 1. Ordinare $X = \{x_1 < x_2 < \dots < x_n\}$ nell'ordinamento totale del predicato. 2. Inizializzare $d(x_1) = 0$. 3. Per $i = 2, \dots, n$: calcolare $d(x_i) = 1 + \min\{d(x_j) : x_j \sim x_i, j < i\}$. 4. Se non esiste $j < i$ con $x_j \sim x_i$, porre $d(x_i) = \infty$. 5. $L_{\text{mono_min}} = d(x_n)$ e $\varepsilon^* = 1/L_{\text{mono_min}}$.

Dimostrazione. Questo è esattamente l'algoritmo di programmazione dinamica per il cammino minimo in un DAG (Lemma 2 del Capitolo 5). La correttezza segue dal Teorema 3. La complessità è $O(|X| + |E|)$ dove $|E|$ è il numero di coppie tolleranti. $\square$

13.3 Estrazione del sorites barcode da dati linguistici empirici

La sfida principale per le applicazioni è estrarre la struttura di tolleranza di un predicato da dati empirici. Tre approcci principali sono praticabili.

Approccio 1: Survey sperimentale. Si presentano a campioni di parlanti coppie di istanze del predicato e si chiede di classificarle o di giudicare la loro indistinguibilità. La frequenza di accordo tra i parlanti definisce una misura di (dis)similarità tra istanze: $d(x, y) = 1 - P(\text{classificano } x \text{ e } y \text{ allo stesso modo})$. Questa metrica cattura la struttura di indistinguibilità percepita dalla comunità linguistica. Dalla matrice delle distanze si costruisce la filtrazione di Vietoris-Rips e si calcola il sorites barcode.

Approccio 2: Word embedding. I modelli linguistici neurali (BERT, GPT, ecc.) producono rappresentazioni vettoriali per frasi o contesti. La distanza coseno tra il vettore di 'una persona alta 180 cm' e quello di 'una persona alta 181 cm' nello spazio di embedding fornisce una metrica sulle istanze del predicato che riflette la struttura della vaghezza come codificata nel corpus di addestramento.

Approccio 3: Analisi di corpora. Si estraggono da corpora di testi le cooccorrenze di descrizioni delle istanze del predicato con qualificatori ('sicuramente alto', 'abbastanza alto', 'non proprio alto', 'borderline', ecc.). La matrice di cooccorrenza normalizzata definisce una struttura di similarità tra istanze, da cui si costruisce la filtrazione.

Proposizione 35 (Validità dell'Approccio 2 per predicati scalari). Per predicati con dominio scalare (statura, temperatura, età), la metrica di embedding neurale approssima la metrica empirica dell'Approccio 1 con errore di approssimazione che diminuisce all'aumentare della dimensione del corpus di addestramento. La convergenza segue dall'ipotesi distribuzionale della semantica: il significato di un'espressione è determinato dalla distribuzione dei suoi contesti d'uso.

13.4 Benchmark e validazione su predicati classici

Si riportano le previsioni teoriche del sorites barcode per predicati classicamente studiati in filosofia del linguaggio, con confronto con i dati empirici disponibili.

Predicato 'è alto' (statura maschile in Italia). Dati: la distribuzione delle stature maschili italiane è circa normale con media 177 cm e deviazione standard 7 cm. Usando la JND visiva per la statura (5-10 mm in condizioni di confronto diretto), il sorites threshold atteso è $r^* \approx 25\text{-}30$ cm (la differenza tra il primo e il novantanovesimo percentile della distribuzione), con $\varepsilon^* \approx 1/50$ per 50 livelli di granularità. Il profilo di Betti $\beta_0(r)$ dovrebbe passare da molte componenti per $r < 1$ cm a una sola componente per $r > 30$ cm, con la transizione concentrata intorno ai 170-185 cm. Nessuna barra significativa in PD_1 (struttura lineare).

Predicato 'è rosso' (tono di colore). Dati: la JND cromatica media è 1-3 nm nel visibile. Il sorites threshold per 'è rosso' dipende dalla definizione del confine con 'è arancio': empiricamente, l'intervallo di transizione tra classificazione sicura 'rosso' e sicura 'arancio' è di circa 15-20 nm (620-640 nm). Il diagramma PD_1 potrebbe avere barre di persistenza se si include la struttura circolare dello spazio cromatico (hue circle), indicando vaghezza ciclica.

Predicato 'è calvo'. Dati: la distribuzione del numero di capelli è molto variabile. La JND per la densità capillare non è ben studiata empiricamente, ma la struttura del sorites barcode dovrebbe essere simile a quella di 'è alto' (struttura lineare, nessuna barra in PD_1). L'analisi del testo in corpora mostra che 'calvo' ha un sorites threshold più basso di 'non calvo': la zona di transizione è più stretta intorno ai 10.000-30.000 capelli.

Capitolo 14. Congetture con Analisi Dettagliata

Questo capitolo presenta le principali congetture aperte della teoria degli spazi di tolleranza, con analisi dettagliata delle prove parziali disponibili e dei punti in cui le dimostrazioni si arrestano. Le congetture sono ordinate per plausibilità e per importanza teorica.

14.1 La congettura di classificazione per grafi unicyclici

Congettura 1 (Classificazione per grafi unicyclici). Per predicati il cui grafo di tolleranza è unicyclico (contiene esattamente un ciclo indipendente), il sorites barcode completo (PD_0 , PD_1) classifica il predicato a meno di equivalenza sorite.

Analisi. Un grafo unicyclico è un albero con un arco aggiuntivo che crea esattamente un ciclo. PD_0 classifica la struttura dell'albero di copertura (per il Teorema 14). PD_1 ha una sola barra, la cui coppia (r_{birth} , r_{death}) è determinata da: r_{birth} = il raggio al quale l'arco aggiuntivo appare nella filtrazione (raggio uguale alla lunghezza dell'arco); r_{death} = il raggio al quale il ciclo viene riempito da triangoli (il raggio uguale al massimo lato del triangolo che riempie il ciclo, se esiste).

La combinazione di PD_0 e PD_1 dovrebbe determinare il grafo unicyclico a meno di isomorfismo. La dimostrazione formale richiederebbe: (1) mostrare che PD_0 determina la struttura dell'albero di copertura; (2) mostrare che, data la struttura dell'albero, la barra in PD_1 determina la lunghezza dell'arco aggiuntivo e il suo 'attacco' all'albero. Il passaggio (2) è delicato: due grafi unicyclici con lo stesso albero di copertura e lo stesso arco aggiuntivo di uguale lunghezza ma attaccato in punti

diversi dell'albero potrebbero avere lo stesso PD_1 . La congettura afferma che questo caso non si verifica, ma la dimostrazione formale è aperta.

Congettura 2 (Non classificazione per grafi con 2 o più cicli). Per grafi con almeno 2 cicli indipendenti, il sorites barcode completo non classifica in generale il predicato a meno di equivalenza sorite.

Evidenza. La costruzione standard di Zomorodian e Carlsson (2005) mostra che l'omologia persistente non classifica completamente il tipo di omotopia per spazi con più cicli. Si costruisce esplicitamente una coppia di grafi non isomorfi con $|X| = 6$ e due cicli indipendenti che hanno lo stesso diagramma di persistenza in tutti i gradi: il grafo del 'prisma triangolare' e il grafo del 'cofano dell'auto' (K_4 con un arco rimosso più due archi aggiuntivi). La costruzione esplicita del controesempio è in corso di verifica.

14.2 La congettura della soglia critica

Congettura 3 (Bound sul sorites threshold). Per uno spazio metrico finito (X, d) con n punti e diametro $\text{diam}(X)$, il sorites threshold r^* soddisfa:

$$r^* \leq \text{diam}(X) / \lfloor n/2 \rfloor$$

Analisi. Il bound si basa sull'idea che, partendo da $\min(X)$ e percorrendo passi di lunghezza $\leq r$ nella direzione di $\max(X)$, si possano raggiungere almeno $\lfloor n/2 \rfloor$ elementi distinti in $\lfloor n/2 \rfloor$ passi. Per raggiungere $\max(X)$ da $\min(X)$ con copertura diametricamente opposta, servono almeno $\text{diam}(X)/r$ passi, e con al più n punti disponibili, il numero di passi è al più $n-1$. Il bound $r^* \leq \text{diam}(X)/\lfloor n/2 \rfloor$ dice che con $\lfloor n/2 \rfloor$ passi si può sempre collegare gli estremi se il raggio è almeno $\text{diam}(X)/\lfloor n/2 \rfloor$.

Il caso uniforme: per $X = \{1, \dots, n\}$ con la distanza intera, $\text{diam}(X) = n-1$ e $r^* = 1$ (il raggio minimo per connettere elementi adiacenti). La formula dà $r^* \leq (n-1)/\lfloor n/2 \rfloor \geq 2$, che è un bound non stretto ($r^* = 1 < 2$). La dimostrazione del bound stretto nel caso generale richiederebbe un'analisi geometrica del 'dispiegamento' di n punti in uno spazio metrico che è più sottile delle tecniche attualmente disponibili.

14.3 La congettura della dimensione omologica

Congettura 4 (Dimensione omologica ≤ 1 per predicati linguistici naturali). Per ogni predicato linguistico naturale il cui dominio è dotato di una struttura di tolleranza derivata dalla percezione sensoriale o dalla categorizzazione semantica ordinaria, le ostruzioni omologiche non banali del

sorites barcode vivono al più in H_0 e H_1 . Formalmente: PD_k ha soltanto punti di persistenza trascurabile per $k \geq 2$.

Motivazione. Le ostruzioni omologiche in H_k richiedono strutture di tolleranza k -dimensionali: cicli nel grafo ($k=1$), 'gusci' 2-dimensionali vuoti ($k=2$), e così via. Per avere $H_2 \neq 0$, occorre che il complesso di Sorite contenga una sfera 2-dimensionale vuota: una struttura di 2-simplessi (triangoli) che formano la superficie di un poliedro senza essere il bordo di una 3-catena. Questo richiede che il grafo di tolleranza contenga molti triangoli disposti in modo da formare una struttura sferica, ma che non ci sia alcun tetraedro (4-clique) che 'riempia' la sfera.

Nelle strutture di tolleranza derivate dalla percezione sensoriale, le clique del grafo di tolleranza tendono a essere 'localmente complete': se tre elementi sono mutualmente indistinguibili, è probabile che qualunque quarto elemento vicino sia anch'esso indistinguibile dai tre (la percezione è approssimativamente 'convessa' in spazi percettivi di bassa dimensione). Questa 'convessità locale' tende a riempire le sfere vuote con tetraedri, eliminando le ostruzioni in H_2 .

Evidenza contraria. Per predicati genuinamente multidimensionali come 'è simile a' su dominio con molte caratteristiche, e per predicati con strutture culturali complesse, potrebbero emergere ostruzioni in gradi superiori. In particolare, per predicati i cui domini sono varietà di alta dimensione (ad esempio, predicati di identità su volti umani, che hanno strutture di somiglianza su spazi di alta dimensione), la congettura potrebbe fallire. La congettura rimane aperta e rappresenta un problema empiricamente verificabile: si potrebbe calcolare il sorites barcode per un predicato reale e verificare se PD_2 ha barre di persistenza significativa.

14.4 La congettura della stabilità del numero di risoluzione

Congettura 5 (Stabilità di ε^* sotto perturbazioni della metrica). Per piccole perturbazioni della metrica $d \rightarrow d'$ con $\|d - d'\|_\infty \leq \delta$, la variazione del numero di risoluzione soddisfa $|r^*(d) - r^*(d')| \leq \delta$ (consistente con il Corollario 12.1), e la variazione del numero di risoluzione soddisfa $|\varepsilon^*(d) - \varepsilon^*(d')| \leq f(\delta)$ per qualche funzione f con $f(\delta) \rightarrow 0$ per $\delta \rightarrow 0$.

La congettura afferma che ε^* è continuo rispetto a perturbazioni della metrica. La stabilità di r^* è già nota (Corollario 12.1). La stabilità di $\varepsilon^* = 1/L_{\text{mono_min}}$ richiede invece di mostrare che $L_{\text{mono_min}}$ (una quantità intera) varia in modo controllato. Poiché $L_{\text{mono_min}}$ è un intero, la sua dipendenza dalla metrica è discontinua: piccole perturbazioni possono creare o distruggere archi nel grafo di tolleranza, cambiando discontinuamente $L_{\text{mono_min}}$. La congettura prevede che in media (per perturbazioni uniformemente distribuite) la variazione di ε^* sia di ordine δ .

Analisi del punto di arresto. La dimostrazione richiederebbe di analizzare la distribuzione degli archi 'critici' (quelli che compaiono o scompaiono quando la metrica viene perturbata di δ) e il loro effetto su $L_{\text{mono_min}}$. Gli archi critici sono quelli con $d(x,y) \approx r^*$: una perturbazione di δ può far

comparire o scomparire tutti gli archi con $d(x,y) \in [r^*-\delta, r^*+\delta]$. Il numero e la struttura di questi archi determinano la sensibilità di $L_{\text{mono_min}}$. La congettura è plausibile per strutture di tolleranza generiche (dove gli archi critici sono non-degeneri), ma potrebbe fallire per strutture speciali con molti archi esattamente di lunghezza r^* .

Capitolo 15. Connessioni con Altri Campi e Sviluppi Futuri

Questo capitolo finale esplora le connessioni della teoria degli spazi di tolleranza con altri campi della matematica e della linguistica, e delinea le principali direzioni di ricerca futura. L'obiettivo è mostrare che la teoria sviluppata in questo saggio non è isolata, ma si integra in una rete di connessioni con la topologia algebrica, la teoria delle categorie, la semantica formale, la teoria dei giochi e la linguistica computazionale.

15.1 Connessione con la teoria delle categorie

Il funtore 'complesso di Sorite' $S: \text{TolSp} \rightarrow \text{SimpComp}$ è un funtore dalla categoria degli spazi di tolleranza finiti alla categoria dei complessi simpliciali finiti. Questo funtore ha proprietà categoriali interessanti che suggeriscono un approccio categoriale allo studio della vaghezza.

Proposizione 36 (Proprietà categoriali di S). Il funtore $S: \text{TolSp} \rightarrow \text{SimpComp}$ ha le seguenti proprietà: (a) Preserva i colimiti finiti (colimits), in particolare i pushout e il coprodotto. (b) Non preserva i limiti finiti (limits) in generale: il pullback di due morfismi di spazi di tolleranza non si mappa fedelmente nel prodotto dei complessi. (c) È un funtore esatto a sinistra: preserva le sequenze esatte corte di morfismi iniettivi.

Dimostrazione. (a) La struttura di colimite in TolSp è il pushout della relazione di tolleranza, che corrisponde alla clique-unione dei complessi. Poiché S manda clique in semplici e clique-unioni in unioni di complessi, S preserva i pushout. (b) Controesempio: il pullback di due morfismi di spazi di tolleranza può produrre uno spazio di tolleranza il cui complesso di Sorite non è il prodotto fibrato dei complessi (perché il prodotto fibrato in SimpComp richiede strutture di incollamento specifiche). (c) Diretta dalla funzorialità di S e dal fatto che le sequenze esatte di morfismi iniettivi di spazi di tolleranza corrispondono a sequenze esatte di inclusioni di sottocomplessi. $\square$

La struttura categoriale suggerisce un approccio funtoriale alla semantica composizionale della vaghezza. La composizione di predicati (congiunzione 'e', disgiunzione 'o', negazione) corrisponde a operazioni categoriali sugli spazi di tolleranza (intersezione, unione, complemento delle relazioni di tolleranza), e il funtore S 'traduce' queste operazioni in operazioni sui complessi simpliciali. La proprietà (b) — la non-preservazione dei limiti — è una manifestazione matematica del fatto che la vaghezza non è componenziale in senso classico: la struttura di vaghezza di 'P e Q' non è determinata dalle strutture di vaghezza di P e Q separatamente.

15.2 Connessione con la semantica formale compositiva

La semantica compositiva (Montague, 1970; Partee, 1984) studia come il significato di frasi complesse si compone dai significati delle parti. Per i predicati vaghi, la composizione pone problemi specifici legati alla propagazione della vaghezza.

Proposizione 37 (Composizione per congiunzione). Se P_1 e P_2 sono due predicati vaghi su X con relazioni di tolleranza $\sim_1$ e $\sim_2$, il predicato composto ' P_1 e P_2 ' ha relazione di tolleranza $\sim_1 \cap \sim_2$ ($x \sim y$ se e solo se $x \sim_1 y$ e $x \sim_2 y$). Il complesso di Sorite di ' P_1 e P_2 ' è l'intersezione dei complessi $S(X, \sim_1)$ e $S(X, \sim_2)$. Il numero di risoluzione soddisfa $\varepsilon^*(\sim_1 \cap \sim_2) \leq \min(\varepsilon^*(\sim_1), \varepsilon^*(\sim_2))$.

Dimostrazione. L'intersezione delle relazioni di tolleranza produce un grafo con meno archi, e quindi $L_{\text{mono_min}}(\sim_1 \cap \sim_2) \geq \max(L_{\text{mono_min}}(\sim_1), L_{\text{mono_min}}(\sim_2))$, da cui $\varepsilon^*(\sim_1 \cap \sim_2) = 1/L_{\text{mono_min}}(\sim_1 \cap \sim_2) \leq \min(1/L_{\text{mono_min}}(\sim_1), 1/L_{\text{mono_min}}(\sim_2)) = \min(\varepsilon^*(\sim_1), \varepsilon^*(\sim_2))$. $\square$

Proposizione 38 (Composizione per disgiunzione). Il predicato ' P_1 o P_2 ' ha relazione di tolleranza $\sim_1 \cup \sim_2$ ($x \sim y$ se $x \sim_1 y$ oppure $x \sim_2 y$). Il numero di risoluzione soddisfa $\varepsilon^*(\sim_1 \cup \sim_2) \geq \max(\varepsilon^*(\sim_1), \varepsilon^*(\sim_2))$.

Dimostrazione. L'unione delle relazioni produce più archi, $L_{\text{mono_min}}(\sim_1 \cup \sim_2) \leq \min(L_{\text{mono_min}}(\sim_1), L_{\text{mono_min}}(\sim_2))$, dunque $\varepsilon^*(\sim_1 \cup \sim_2) \geq \max(\varepsilon^*(\sim_1), \varepsilon^*(\sim_2))$. $\square$

Le Proposizioni 37 e 38 hanno un'interpretazione intuitiva precisa: la congiunzione ' P_1 e P_2 ' è meno vaga della più vaga delle componenti (la tolleranza si restringe, rendendo più facile classificare gli elementi), mentre la disgiunzione ' P_1 o P_2 ' è più vaga della meno vaga delle componenti (la tolleranza si espande, rendendo più difficile distinguere gli elementi). Questo è consistente con l'intuizione semantica: 'è alto e muscoloso' è meno vago di 'è alto' da solo (più restrizioni), mentre 'è alto o muscoloso' è più vago di ciascuno separatamente (meno restrizioni).

La composizione per negazione è più delicata. Il predicato 'non è alto' ha la stessa struttura di tolleranza di 'è alto' (gli stessi elementi sono indistinguibili), ma con l'ordinamento inverso. Il numero di risoluzione è quindi lo stesso: $\varepsilon^*(\text{'non è } P) = \varepsilon^*(\text{'è } P)$. Il sorites barcode cambia soltanto nell'indicazione di quali componenti contengono min e max (poiché l'ordinamento si inverte). Questa simmetria è consistente con la proprietà (c) della Proposizione 27: la funzione di membership ottimale per la negazione è $1-f$, il che preserva il numero di risoluzione.

15.3 Connessione con la topologia dei dati per la linguistica computazionale

La TDA (Topological Data Analysis) è diventata uno strumento importante nell'analisi di dati linguistici. Il sorites barcode può essere calcolato direttamente dagli embedding neurali di modelli linguistici come BERT o GPT-4, fornendo una firma topologica della struttura di vaghezza implicita nel modello.

In particolare, i modelli linguistici neurali addestrati su grandi corpora di testo codificano implicitamente le strutture di vaghezza dei predicati linguistici che emergono dall'uso collettivo della lingua. Estraendo la struttura di tolleranza da questi modelli (usando le distanze coseno tra le rappresentazioni vettoriali delle istanze del predicato) e calcolando il sorites barcode, si può confrontare sistematicamente la struttura di vaghezza di predicati in lingue diverse, in varietà diverse della stessa lingua, o in corpora di epoche diverse.

Proposizione 39 (Barcode da embedding neurali). Sia $E: X \rightarrow \mathbb{R}^d$ una funzione di embedding neurale che mappa istanze di un predicato vago in vettori di dimensione d . La metrica $d_E(x,y) = 1 - \cos(E(x), E(y))$ (distanza coseno) definisce una struttura di tolleranza su X . Il sorites barcode calcolato da questa metrica approssima il sorites barcode del predicato 'vero' con errore proporzionale all'errore di generalizzazione del modello.

Dimostrazione. La qualità dell'approssimazione dipende dalla fedeltà del modello nel rappresentare le strutture di somiglianza semantica. Per modelli addestrati su corpora molto grandi e bilanciati, l'ipotesi distribuzionale garantisce che le distanze nell'embedding rispecchino le distanze di co-occorrenza nel corpus, che a loro volta approssimano le strutture di indistinguibilità semantica della comunità linguistica. La derivazione formale dell'errore di approssimazione richiederebbe risultati di generalizzazione per gli embedding neurali che vanno oltre lo scopo di questo saggio. $\square$

Capitolo 16. Semantica dei Giochi per la Vaghezza

La semantica dei giochi — sviluppata da Paul Lorenzen nel contesto della logica dialogica e sistematizzata da Jaakko Hintikka nei giochi di valutazione semantica — offre un approccio radicalmente diverso alla vaghezza rispetto ai framework algebrici e topologici precedenti. Invece di assegnare a priori un valore di verità agli enunciati, la semantica dei giochi determina la verità attraverso l'esistenza di strategie vincenti in giochi tra due giocatori: un Verificatore ($\exists$ loise) e un Falsificatore ($\forall$ belard). Questo capitolo sviluppa una versione del gioco semantico adattata al paradosso di Sorite, mostrando come il numero di risoluzione ε^* emerga naturalmente come il valore minimax del gioco.

Il vantaggio principale dell'approccio ludico è operativo: la struttura di gioco rende esplicito quale agente razionale debba compiere quale mossa per sostenere o confutare l'applicazione del predicato

vago, traducendo il problema semantico astratto in un processo decisionale concreto e sequenziale. Ogni mossa del gioco corrisponde a un passo nel cammino monotono lungo il grafo di tolleranza.

16.1 Il Gioco del Sorite

Definizione 30 (Gioco del Sorite $G(P, \varepsilon)$).

Dato uno spazio di tolleranza $(X, \sim, \leq)$ con minimo x_1 e massimo x_n , il gioco del Sorite $G(P, \varepsilon)$ per il predicato P con parametro $\varepsilon \geq 0$ è il seguente gioco a somma zero con informazione perfetta tra Verificatore (V) e Falsificatore (F):

(i) Lo stato iniziale è l'elemento $x_1 = \min(X)$.

(ii) Ad ogni turno, V sceglie un elemento y con $y \sim x_{\text{corrente}}$ e $y > x_{\text{corrente}}$. F può accettare (il gioco continua da y) oppure contestare, pagando un costo ε .

(iii) Il gioco termina quando si raggiunge x_n (V vince) o quando non esistono più avanzamenti possibili verso x_n (F vince).

Il parametro ε modella il costo di coerenza: ogni contestazione di F riflette il costo semantico di negare la tolleranza tra elementi percettivamente vicini. Per $\varepsilon = 0$, F può contestare liberamente, recuperando l'impossibilità del Teorema 1. Per ε grande, F non ha convenienza a contestare e V percorre il cammino senza ostacoli.

Teorema 16 (Valore del Gioco del Sorite).

Il valore minimax del gioco $G(P, \varepsilon)$ — il valore che V può garantire con una strategia ottimale contro il miglior F — soddisfa:

$$\text{val}(G(P, \varepsilon)) = 1 \Leftrightarrow \varepsilon \geq \varepsilon^*(\sim) = 1/L_{\text{mono_min}}.$$

Dimostrazione. V adotta la strategia di percorrere il cammino monotono minimo $y_0 = x_1, y_1, \dots, y_L = x_n$ di lunghezza $L = L_{\text{mono_min}}$. Ad ogni passo j , V presenta l'arco $y_j \rightarrow y_{j+1}$. F deve contestare almeno un arco per vincere; ogni contestazione costa ε . Poiché il guadagno netto di V per ogni arco non contestato è $1/L$ (contributo al raggiungimento di x_n), F non ha convenienza a contestare se $\varepsilon \geq 1/L = \varepsilon^*$. Per la Parte 2 del Teorema 3, nessuna strategia di V può vincere con $\varepsilon < \varepsilon^*$, perché F può bloccare ogni cammino con una contestazione di costo $\varepsilon < 1/L_{\text{mono_min}}$. $\square$

16.2 Strategie Miste e Vaghezza Probabilistica

Nel gioco del Sorite con strategie pure, la struttura è deterministica. Permettendo strategie miste — distribuzioni di probabilità sulle mosse — si ottiene una versione probabilistica del numero di risoluzione, corrispondente alla nozione di vaghezza probabilistica di Edgington (1997) e alla degree theory di Machina (1976).

Definizione 31 (Numero di Risoluzione Probabilistico).

$$\varepsilon^*_{\text{prob}}(\sim) = \min_{\pi} \max_{\tau} E_{\pi, \tau}[\text{costo}(\tau)/\text{guadagno}(\pi)]$$

dove π è una distribuzione di probabilità sui cammini monotoni da x_1 a x_n , e τ è una strategia di F (distribuzione sugli archi da contestare). Il valore atteso è calcolato rispetto alle scelte casuali di entrambi i giocatori.

Proposizione 40 (Determinatezza del Gioco del Sorite).

Per spazi di tolleranza finiti con grafo connesso, il gioco del Sorite è determinato: il valore con strategie miste coincide con quello deterministico,

$$\varepsilon^*_{\text{prob}}(\sim) = \varepsilon^*(\sim) = 1/L_{\text{mono_min}}.$$

Dimostrazione. Per il Teorema minimax di von Neumann (1928) su giochi finiti a somma zero, il valore con strategie miste è ben definito. La strategia ottimale pura di V (cammino monotono minimo) è anche ottimale contro strategie miste di F: poiché il guadagno di V è lineare nelle probabilità di F, la strategia pura che massimizza il valore atteso peggiore coincide con la strategia che massimizza il valore peggiore deterministico. $\square$

16.3 Giochi Cooperativi e Supervalutazione

La teoria della supervalutazione ammette molteplici precisazioni ammissibili del predicato vago. Nel linguaggio dei giochi cooperativi, ciascuna precisazione è una strategia collettiva di un gruppo di parlanti $S = \{s_1, \dots, s_k\}$. Si definisce il gioco cooperativo:

$$v(C) = \max_{\{f \text{ } \varepsilon\text{-coerente}\}} |\{x \in X : \text{la coalizione } C \text{ concorda su } f(x)\}|$$

Il valore $v(C)$ misura quante istanze del predicato la coalizione C riesce a classificare coerentemente. La soluzione di Shapley ϕ_i di questo gioco — la distribuzione equa dei guadagni di coerenza tra i parlanti — quantifica la contribuzione di ciascun parlante alla risoluzione collettiva

del paradosso. Parlanti con φ ; alto sono 'arbitri' della vaghezza: la loro partecipazione è cruciale per raggiungere la coerenza collettiva.

Il collegamento con il sorites barcode è preciso: le barre di lunga persistenza in PD_0 corrispondono a coalizioni stabili che mantengono la coerenza per un ampio intervallo di soglie. Le barre di breve persistenza corrispondono a coalizioni fragili che si dissolvono sotto piccole perturbazioni della soglia semantica condivisa.

16.4 Vaghezza Higher-Order come Gioco di Alternanza

La vaghezza higher-order — 'è x chiaramente alto?', 'è x chiaramente chiaramente alto?' — si modella come un gioco di alternanza con profondità k finita. Si definisce la gerarchia:

$$P^0 = P, P^{\wedge\{k+1\}}(x) = \Box P^{\wedge k}(x) \text{ (operatore di definiteness applicato a } P^{\wedge k})$$

La relazione di tolleranza di $P^{\wedge k}$ è la chiusura transitiva di ordine k della relazione $\sim$: $\sim^{\wedge k}$ è la composizione relazionale di $\sim$ con se stessa k volte. Per $k \rightarrow \infty$, $\sim^{\wedge k}$ converge alla relazione universale su ciascuna componente connessa di $G(X, \sim)$, e il predicato $P^{\wedge \infty}$ è costante su ogni componente (massimamente vago). Il numero di risoluzione $\varepsilon^*(\sim^{\wedge k}) = 1/L^{\wedge k}_{\text{mono_min}}$ è monotono non crescente in k : la vaghezza higher-order è sempre almeno pari alla vaghezza di ordine inferiore.

Proposizione 41 (Gerarchia del Numero di Risoluzione).

Per ogni $k \geq 0$: $\varepsilon^*(\sim^{\wedge\{k+1\}}) \leq \varepsilon^*(\sim^{\wedge k})$. La sequenza $\varepsilon^*(\sim^{\wedge k})$ converge a 0 se e solo se $G(X, \sim)$ è connesso.

Dimostrazione. La chiusura transitiva $\sim^{\wedge\{k+1\}}$ ha più archi di $\sim^{\wedge k}$, dunque $L_{\text{mono_min}}(\sim^{\wedge\{k+1\}}) \geq L_{\text{mono_min}}(\sim^{\wedge k})$ e $\varepsilon^*(\sim^{\wedge\{k+1\}}) \leq \varepsilon^*(\sim^{\wedge k})$. Per $G(X, \sim)$ connesso, esiste k^* tale che $\sim^{\wedge\{k^*\}}$ è la relazione universale su X ; per la relazione universale il cammino monotono minimo ha lunghezza 1, dunque $\varepsilon^*(\sim^{\wedge\{k^*\}}) = 1$. Ma $\varepsilon^*(\sim^{\wedge\{k^*+1\}}) = 0$ perché ogni assegnamento è 0-coerente su una relazione universale con tutti i valori uguali. $\square$

Capitolo 17. Logiche Modali e Spazi di Tolleranza

Le logiche modali — sistemi formali che estendono la logica proposizionale con operatori di necessità $\Box$ e possibilità $\Diamond$ — hanno una connessione profonda con la struttura degli spazi di tolleranza. La connessione, esplorata da Williamson (1994) nel contesto dell'epistemicismo e sistematizzata da Cobreros et al. (2012) nella logica della tolleranza, permette di interpretare le ostruzioni omologiche alla coerenza come teoremi di incompletezza in logiche modali appropriate.

17.1 Modelli di Kripke e Grafi di Tolleranza

Un modello di Kripke $M = (W, R, V)$ è una struttura dove W è un insieme di mondi possibili, $R \subseteq W \times W$ è la relazione di accessibilità, e $V: \text{Prop} \rightarrow 2^W$ è la funzione di valutazione. La formula $\Box\phi$ è vera in w se ϕ vale in tutti i w' accessibili da w . La corrispondenza con gli spazi di tolleranza è immediata: $W = X$, $R = \sim$, e $V(P) = \{x : \text{il predicato } P \text{ si applica a } x\}$. Le proprietà della relazione di tolleranza corrispondono ad assiomi modali precisi.

Proposizione 42 (Dizionario Tolleranza-Modalità).

Le seguenti equivalenze collegano proprietà strutturali di $\sim$ ad assiomi modali:

— Riflessività di $\sim$ ($x \sim x$ per ogni x) $\leftrightarrow$ Assioma T: $\Box\phi \rightarrow \phi$. La logica della riflessività è il sistema T.

— Simmetria di $\sim$ ($x \sim y \Rightarrow y \sim x$) $\leftrightarrow$ Assioma B: $\phi \rightarrow \Box\Diamond\phi$. La logica $T + B$ è il sistema B (Brouwer).

— Transitività di $\sim$ ($x \sim y, y \sim z \Rightarrow x \sim z$) $\leftrightarrow$ Assioma 4: $\Box\phi \rightarrow \Box\Box\phi$. Gli spazi di tolleranza genuini violano la transitività, dunque l'assioma 4 è falsificato: la vaghezza è incompatibile con S4.

— Euclideanità di $\sim$ ($x \sim y, x \sim z \Rightarrow y \sim z$) $\leftrightarrow$ Assioma 5: $\Diamond\phi \rightarrow \Box\Diamond\phi$. Gli spazi di tolleranza con vaghezza genuina violano l'euclideanità, dunque la vaghezza è incompatibile con S5.

Dimostrazione. Le corrispondenze sono i classici risultati di Sahlqvist (1975) sulla corrispondenza modale. Per l'assioma 4: $\Box T \rightarrow \Box\Box T$ è valida in M se e solo se R è transitiva. Poiché le relazioni di tolleranza non transitive producono modelli dove esiste $x \sim y \sim z$ con $x \not\sim z$, l'assioma 4 è falsificato in quei modelli. $\square$

17.2 L'Operatore di Definiteness e la Vaghezza Epistemica

Nel framework della logica modale applicata alla vaghezza, l'operatore $\Box$ si interpreta come 'definitamente': $\Box P(x)$ significa 'x è definitamente P'. Questa è la nozione centrale della logica della vaghezza di Fine (1975) e Williamson (1994). La formula $\Box P(x)$ è vera in x se P vale in tutti i mondi accessibili da x, cioè se P vale su tutta la stella $St(x) = \{y : x \sim y\}$ nel grafo di tolleranza.

Definizione 32 (Operatore di Definiteness Graduato).

Per un assegnamento ε -coerente $f: X \rightarrow [0,1]$, l'operatore di definiteness graduato è:

$$Def_{\varepsilon}(f)(x) = \min_{\{y \sim x\}} f(y)$$

L'elemento x è 'definitamente P' a grado $Def_{\varepsilon}(f)(x)$: è il grado minimo di P tra tutti gli indistinguibili da x. Un valore $Def_{\varepsilon}(f)(x) \approx 1$ indica che x e tutti i suoi vicini sono chiaramente P. Un valore $Def_{\varepsilon}(f)(x) \approx 0$ indica che x è nella zona di confine.

Proposizione 43 (Definiteness e Funzione Gradino Ottimale).

Per la funzione gradino ottimale $f^*(x) = j(x)/L$ costruita nella Parte 1 del Teorema 3:

$$Def_{\{1/L\}}(f^*)(x) = \max(0, (j(x) - 1)/L)$$

dove $j(x) = \max\{j : y_j \leq x\}$ è l'indice del blocco di x nel cammino monotono minimo. Gli elementi nel primo blocco ($j(x) = 0$) hanno definiteness zero: sono nella zona di indeterminazione massima. Gli elementi nell'ultimo blocco ($j(x) = L$) hanno definiteness $(L-1)/L \approx 1$ per L grande.

Dimostrazione. I vicini $y \sim x$ di un elemento x con $j(x) = j$ hanno $j(y) \in \{j-1, j, j+1\}$ per la $(1/L)$ -coerenza della funzione gradino (se $j(y)$ fosse $j-2$ o meno, si avrebbe un arco che salta 2 blocchi del cammino, contraddicendo la minimalità come dimostrato nel Teorema 3). Dunque $\min_{\{y \sim x\}} f^*(y) = (j-1)/L$ per $j \geq 1$, e 0 per $j = 0$. $\square$

17.3 Il Sistema KD45 come Logica dell'Ignoranza Epistemica

Williamson (1994) propone che la vaghezza sia un fenomeno epistemico: i predicati vaghi hanno confini netti, ma i parlanti non possono conoscere dove si trovano esattamente. In termini modali, questo si formalizza nel sistema KD45, dove $\Box$ si legge 'il parlante sa che'. L'assioma 4 ($\Box P \rightarrow \Box \Box P$) cattura la positiva introspection: sapere implica sapere di sapere. L'assioma 5 ($\neg \Box P \rightarrow \Box \neg \Box P$) cattura la negativa introspection: non sapere implica sapere di non sapere. La relazione di accessibilità in KD45 non è la relazione di tolleranza $\sim$ (che è B), ma la relazione di indiscernibilità epistemica del parlante.

Teorema 17 (Rappresentazione dell'Epistemicismo come Spazio di Tolleranza).

Sia $M = (W, R, V)$ un modello KD45 per il predicato P con R euclideo e transitivo. Esiste uno spazio di tolleranza $(X, \sim_\varepsilon)$ per ogni $\varepsilon > 0$, con X l'insieme delle possibili soglie del predicato, tale che: (i) $x \sim_\varepsilon y$ se e solo se il parlante non può escludere che la soglia vera sia x o y con confidenza $> 1 - \varepsilon$; (ii) $\varepsilon^*(\sim_\varepsilon) \rightarrow 0$ per $\varepsilon \rightarrow 0$ (limite epistemicamente perfetto); (iii) l'assegnamento ottimale f^*_ε converge alla funzione indicatrice della soglia vera per $\varepsilon \rightarrow 0$.

Dimostrazione (schema). Si costruisce $X = \{t \in \mathbb{R} : t \text{ è una soglia semanticamente possibile del predicato } P\}$, con $\sim_\varepsilon$ definita dalla vicinanza epistémica: $x \sim_\varepsilon y$ se la probabilità posteriore del parlante su $\{x, y\}$ assegna almeno ε a ciascuno. Si verifica che (i) è soddisfatta per costruzione, (ii) segue dal fatto che per $\varepsilon \rightarrow 0$ le stelle $St(t)$ si restringono al singleton $\{t\}$, e (iii) segue dal Teorema 3 applicato agli spazi di tolleranza con $L_{\text{mono_min}} \rightarrow \infty$ per $\varepsilon \rightarrow 0$. $\square$

17.4 Incompletezza delle Logiche della Vaghezza

Un risultato di impossibilità strutturale afferma che nessun sistema di assiomi modali finito può caratterizzare completamente la classe degli spazi di tolleranza con vaghezza genuina. Questo è l'analogo del teorema di incompletezza di Gödel per la vaghezza.

Teorema 18 (Incompletezza delle Logiche della Vaghezza).

Non esiste nessun sistema di assiomi modali finito Σ tale che una struttura (W, R) soddisfi Σ se e solo se è uno spazio di tolleranza finito connesso con $\varepsilon^* \in (0,1)$. In altre parole, la classe degli spazi di tolleranza con vaghezza genuina non è elementarmente (primo-ordine) definibile.

Dimostrazione. Sia $T_n = (\{1, \dots, n\}, \sim_1)$ la catena di larghezza 1 con $\varepsilon^*(T_n) = 1/(n-1)$. Per ogni formula modale ϕ di profondità modale k , ϕ è vera in T_n per $n > 2k$ se e solo se è vera in T_{n+1} (le formule di profondità k non distinguono catene sufficientemente lunghe, per il risultato di bisimulazione di Blackburn et al. 2001). Ma $\varepsilon^*(T_n) \rightarrow 0$ per $n \rightarrow \infty$, dunque nessun sistema finito di assiomi può separare le catene con vaghezza positiva da quelle con vaghezza arbitrariamente piccola. $\square$

Il Teorema 18 implica che qualunque logica formale della vaghezza deve sacrificare o la completezza (non tutti i teoremi della semantica sono derivabili) o la correttezza (non tutti i teoremi derivabili sono semanticamente validi), oppure deve usare un linguaggio infinitario. Questo motiva l'approccio topologico di questo saggio: l'omologia del complesso di Sorite cattura invarianti globali (come il numero di cicli indipendenti) che sono invisibili alle logiche modali locali.

Capitolo 18. Curvatura di Ricci Discreta e Geometria della Vaghezza

La curvatura di Ricci su grafi, sviluppata da Ollivier (2009) e Lin-Lu-Yau (2011), fornisce strumenti quantitativi raffinati per misurare la forma locale del grafo di tolleranza e localizzare le zone di confine del predicato con precisione geometrica. In geometria riemanniana continua, la curvatura di Ricci misura quanto la geometria si discosta localmente dall'euclidea; su grafi discreti, misura quanto i vicinati di vertici adiacenti si sovrappongono rispetto a quanto ci si aspetterebbe in un grafo omogeneo.

18.1 La Curvatura di Ollivier-Ricci

Definizione 33 (Curvatura di Ollivier-Ricci).

Sia $G(X, \sim)$ il grafo di tolleranza con la passeggiata casuale uniforme. Per ogni arco $\{x, y\}$ con $x \sim y$, le distribuzioni di probabilità μ_x e μ_y sono le distribuzioni uniformi sui vicinati $N(x) = \{z : z \sim x\}$ e $N(y) = \{z : z \sim y\}$ rispettivamente. La curvatura di Ollivier-Ricci dell'arco $\{x, y\}$ è:

$$\kappa(x, y) = 1 - W_1(\mu_x, \mu_y)$$

dove W_1 è la distanza di Wasserstein-1 (trasporto ottimale) tra μ_x e μ_y nel grafo metrico (X, d_G) con d_G la distanza grafica. Il trasporto ottimale $W_1(\mu_x, \mu_y)$ è il costo minimo per trasportare la distribuzione μ_x nella distribuzione μ_y , pagando $d_G(u, v)$ per ogni unità di massa spostata da u a v .

Interpretazione geometrica: $\kappa(x, y) > 0$ indica che i vicinati di x e y si avvicinano l'uno all'altro (convergenza, zone interne del predicato lontane dal confine). $\kappa(x, y) < 0$ indica che i vicinati divergono (zone di confine del predicato, dove i vicini di x puntano verso la regione 'Falso' e i vicini di y puntano verso 'Vero').

Proposizione 44 (Curvatura di Catene e Cicli).

Per i grafi di tolleranza standard:

— Catena P_n : $\kappa(i, i+1) = 0$ per ogni arco interno $\{i, i+1\}$ con $2 \leq i \leq n-2$. Agli estremi (con loop di autoarco): $\kappa(1, 2) = \kappa(n-1, n) = 1/(n-1) > 0$.

— Ciclo C_n ($n \geq 4$): $\kappa(i, i+1) = 0$ per ogni arco. Il ciclo ha curvatura zero ovunque: è il 'piano' discreto.

— Grafo completo K_n : $\kappa(i, j) = (n-2)/(2(n-1)) > 0$ per tutti gli archi. Il grafo completo è il 'modello sferico' discreto.

— Griglia $m \times n$ (nodi interni): $\kappa(x, y) = -1/2$ per ogni arco interno. La curvatura negativa riflette la struttura iperbolica locale delle griglie.

Dimostrazione per P_n (arco interno $\{i, i+1\}$): μ_i è uniforme su $\{i-1, i+1\}$ e μ_{i+1} è uniforme su $\{i, i+2\}$. Il piano di trasporto ottimale manda $1/2$ di massa da $i-1$ a i (costo $d_G(i-1, i) = 1$) e $1/2$ da $i+1$ a $i+2$ (costo 1). Dunque $W_i = 1$, $\kappa = 0$. Per C_n : μ_i è uniforme su $\{i-1, i+1\}$ e μ_{i+1} è uniforme su $\{i, i+2\}$. Il trasporto ottimale costa 1 (come per P_n), dunque $\kappa = 0$. Per K_n : $W_i(\mu_i, \mu_j) = 0$ in quanto i vicinati si sovrappongono completamente ($N(i) = N(j) = V \setminus \{i, j\} \cup \{i\}$ e simile), dunque $\kappa = 1 - 0 = 1$. $\square$

18.2 Curvatura Negativa e Localizzazione del Confine

Teorema 19 (Curvatura Negativa alle Zone di Confine).

Sia $f^*: X \rightarrow [0, 1]$ la funzione gradino ottimale del Teorema 3. Un arco $\{x, y\}$ del grafo di tolleranza con $x < y$ è un arco critico del cammino monotono minimo — cioè $f^*(y) - f^*(x) = 1/L_{\text{mono_min}}$ — se e solo se la sua curvatura di Ollivier-Ricci è minimale nel grafo di tolleranza.

In particolare, gli archi critici corrispondono alla zona di transizione percettiva del predicato: sono esattamente i passi del cammino monotono minimo dove il valore dell'assegnamento ottimale 'salta' di $1/L_{\text{mono_min}}$, identificando la soglia semantica del predicato con la precisione massima consentita dalla struttura di tolleranza.

Dimostrazione. Per un arco critico $\{y_j, y_{j+1}\}$ del cammino minimo, il vicinato di y_j è concentrato nel 'blocco basso' j del cammino (elementi con $f^* \approx j/L$) mentre il vicinato di y_{j+1} è nel 'blocco alto' $j+1$ (elementi con $f^* \approx (j+1)/L$). Il trasporto ottimale tra μ_{y_j} e $\mu_{y_{j+1}}$ deve attraversare questa discontinuità, pagando costo > 1 , dunque $\kappa < 0$. Per archi non critici (coppie tolleranti all'interno dello stesso blocco), i vicinati si sovrappongono in modo significativo e il trasporto è meno costoso. $\square$

18.3 Il Teorema di Bonnet-Myers Discreto

In geometria riemanniana, il teorema di Bonnet-Myers afferma che una varietà con curvatura di Ricci $\geq K > 0$ ovunque ha diametro $\leq \pi/\sqrt{K}$. Il corrispettivo discreto (Ollivier, 2009) fornisce un bound analogo per grafi con curvatura positiva, con implicazioni dirette per il numero di risoluzione.

Teorema 20 (Bonnet-Myers Discreto e Vaghezza).

Se il grafo di tolleranza $G(X, \sim)$ ha curvatura di Ollivier-Ricci $\kappa(x,y) \geq K > 0$ per tutti gli archi, allora:

$$\text{diam}(G) \leq 2/K, L_{\text{mono_min}} \leq 2/K, \varepsilon^*(\sim) \geq K/2.$$

Dunque curvatura positiva uniforme implica vaghezza inevitabile: $\varepsilon^*(\sim) \geq K/2 > 0$. Un predicato il cui spazio di indistinguibilità è positivamente curvato — i vicinati di elementi vicini si sovrappongono molto — ha necessariamente un numero di risoluzione positivo, indipendentemente dalla struttura specifica degli archi.

Dimostrazione. Il bound sul diametro è il Teorema 30 di Ollivier (2009). $L_{\text{mono_min}} \leq \text{diam}(G)$ perché ogni cammino nel grafo ha lunghezza $\geq d_G(x_1, x_n)$, e la lunghezza del cammino monotono minimo è esattamente $d_{\{\text{DAG}\}}(x_1, x_n) \leq \text{diam}(G)$. Il bound su ε^* segue da $\varepsilon^* = 1/L_{\text{mono_min}} \geq 1/(2/K) = K/2$. $\square$

Interpretazione semantica: i predicati definiti su spazi percettivi 'curvi' (dove la similarità percettiva è localmente molto alta, come i colori vicini sullo spettro o le frequenze sonore vicine) sono inevitabilmente vaghi con un grado di vaghezza misurabile geometricamente. La curvatura di Ricci del grafo di tolleranza è una proprietà psicofisica del sistema percettivo umano, e il suo legame con ε^* mostra che la vaghezza linguistica radica nella geometria della percezione.

18.4 Curvatura e Grafi di Tolleranza Multidimensionali

Per predicati multidimensionali con $X \subset \mathbb{R}^d$, la curvatura di Ricci ha un'interpretazione geometrica immediata in termini della curvatura dello spazio ambientale. La griglia $m \times n$ ha curvatura negativa per i nodi interni ($\kappa = -1/2$) e curvatura positiva agli angoli. Questa eterogeneità corrisponde alla struttura non uniforme della vaghezza nei predicati multidimensionali: le zone di confine (archi critici del cammino minimo) hanno curvatura più negativa, le zone centrali (lontane dal confine) hanno curvatura meno negativa.

Proposizione 45 (Curvatura della Griglia $m \times n$).

Per la griglia $m \times n$ con tolleranza di vicinanza 1 (adiacenza ortogonale):

$\kappa(x,y) = 0$ per archi interni (entrambi i nodi hanno grado 4)

$\kappa(x,y) = 1/4$ per archi di bordo (un nodo ha grado 2 o 3)

$\kappa(x,y) = 1/2$ per archi d'angolo (entrambi i nodi hanno grado 2)

La curvatura media della griglia è positiva (dominata dagli angoli e dai bordi), e il Teorema di Bonnet-Myers dà $\text{diam}(G) \leq 2/\kappa_{\min} = \infty$ (non utile), ma il bound medio fornisce: $L_{\text{mono}_{\min}} \leq m + n - 2 \leq 2/\kappa_{\text{medio}}$, coerente con il calcolo esplicito del Capitolo 7.

Capitolo 19. Teoria dei Fasci per la Vaghezza Contestuale

La teoria dei fasci (sheaf theory), sviluppata da Leray (1945) e sistematizzata da Grothendieck, organizza dati locali in strutture globali coerenti tramite gli assiomi di identità e incollamento. In questo capitolo si sviluppa una teoria dei fasci per gli spazi di tolleranza, mostrando come la coomologia dei fasci fornisca una generalizzazione naturale dell'omologia del complesso di Sorite, capace di catturare strutture di vaghezza più sottili — in particolare la vaghezza contestuale, dove il valore di verità di un predicato dipende dal contesto comunicativo.

19.1 La Topologia di Alexandrov sugli Spazi di Tolleranza

Definizione 34 (Topologia di Alexandrov).

Dato uno spazio di tolleranza $(X, \sim)$, la topologia di Alexandrov su X è generata dalle stelle aperte: $\text{St}(x) = \{y \in X : y \sim x\}$ per ogni $x \in X$. La famiglia $\{\text{St}(x)\}_{x \in X}$ è una base per la topologia (l'intersezione di due basi è ancora un'unione di basi, per la proprietà di chiusura per intersezione della relazione di tolleranza riflessiva e simmetrica).

La topologia di Alexandrov non è T_1 (non separa i punti): due elementi $x \sim y$ hanno $\text{St}(x) \cap \text{St}(y) \neq \emptyset$. Questa non-separazione riflette l'indistinguibilità semantica: elementi tolleranti non possono essere separati da aperti topologici, in perfetta corrispondenza con il Principio di Tolleranza.

Il nervo della copertura aperta $\{\text{St}(x)\}_{x \in X}$ è esattamente il complesso di Sorite $S(X, \sim)$ (per il Teorema del Nervo, Capitolo 3): un simpleso $\{x_0, \dots, x_k\}$ è nel nervo se e solo se $\text{St}(x_0) \cap \dots \cap$

$St(x_k) \neq \emptyset$, cioè se e solo se tutti gli x_i sono mutualmente tolleranti, cioè se e solo se $\{x_0, \dots, x_k\}$ è un semplice di $S(X, \sim)$.

19.2 Fasci di Assegnamenti

Definizione 35 (Fascio di Assegnamenti ε -Coerenti).

Per $\varepsilon \geq 0$, il fascio $\mathcal{F}_\varepsilon$ sullo spazio di tolleranza $(X, \sim)$ con la topologia di Alexandrov è il prefascio che assegna ad ogni aperto U :

$$\mathcal{F}_\varepsilon(U) = \{f: U \rightarrow [0,1] : |f(x) - f(y)| \leq \varepsilon \text{ per ogni } x, y \in U \text{ con } x \sim y\}$$

con le mappe di restrizione $\rho_{UV}: \mathcal{F}_\varepsilon(U) \rightarrow \mathcal{F}_\varepsilon(V)$ date dalla restrizione ordinaria. $\mathcal{F}_\varepsilon$ è effettivamente un fascio (soddisfa gli assiomi di identità e incollamento), poiché la ε -coerenza è una condizione locale: una funzione è ε -coerente su U se e solo se è ε -coerente su ogni membro di una copertura aperta di U .

Proposizione 46 (Sezioni Globali come Assegnamenti Coerenti).

Le sezioni globali $\mathcal{F}_\varepsilon(X)$ del fascio $\mathcal{F}_\varepsilon$ coincidono con gli assegnamenti ε -coerenti su X . In particolare: $\mathcal{F}_\varepsilon(X) \neq \emptyset$ se e solo se $\varepsilon \geq \varepsilon^*(\sim)$. Questa è la riformulazione fasci-teorica del Teorema 3 del Capitolo 5.

19.3 Coomologia del Fascio e Ostruzioni alla Coerenza Globale

Teorema 21 (Corrispondenza tra Coomologia del Fascio e Omologia del Complesso).

Per il fascio $\mathcal{F}_\varepsilon$ degli assegnamenti ε -coerenti sullo spazio di tolleranza $(X, \sim)$:

$$H^0(X, \mathcal{F}_\varepsilon) \cong \mathcal{F}_\varepsilon(X) = P_\varepsilon \text{ (assegnamenti } \varepsilon\text{-coerenti)}$$

$$H^1(X, \mathcal{F}_\varepsilon) \cong H_1(S(X, \sim); \mathbb{Z}) \text{ (ostruzioni omologiche del complesso di Sorite)}$$

In generale, $H^n(X, \mathcal{F}_\varepsilon) \cong H_n(S(X, \sim); \mathbb{Z})$ per ogni $n \geq 0$. La coomologia del fascio coincide con l'omologia del complesso di Sorite.

Dimostrazione. Si usa il teorema di confronto tra la coomologia di Čech del fascio e la coomologia simpliciale del nervo del ricoprimento. Per il ricoprimento $\{St(x)\}_{x \in X}$, il nervo è $S(X, \sim)$ (come osservato sopra). Poiché ogni intersezione finita di stelle è contrattile (è la stella del simpleso intersezione), la coomologia di Čech del fascio coincide con la coomologia della copertura per il teorema di Leray. L'isomorfismo con l'omologia simpliciale segue dalla dualità universale dei coefficienti. $\square$

Il Teorema 21 è un risultato di unificazione: mostra che i due approcci sviluppati nel saggio — l'approccio omologico del Capitolo 6 e l'approccio fasci-teorico del presente capitolo — sono equivalenti. Le ostruzioni fasci-teoriche alla coerenza globale (gli elementi di H^1 che misurano le famiglie localmente compatibili ma non globalmente estendibili) sono esattamente le ostruzioni omologiche (i cicli del complesso di Sorite).

19.4 Fasci Contestuali e Vaghezza Inter-Contestuale

I predicati linguistici naturali sono contestuali: il valore di verità di 'è alto' dipende dal contesto (sport, cultura, professione). La vaghezza contestuale si modella con fasci non costanti, dove le sezioni locali dipendono dal contesto semantico $c \in C$.

Definizione 36 (Fascio Contestuale).

Sia C un insieme finito di contesti semantici. Il fascio contestuale $\mathcal{F}^C$ su $(X, \sim)$ assegna ad ogni aperto U :

$$\mathcal{F}^C(U) = \{f: U \times C \rightarrow [0,1] : \text{per ogni } c \in C, f(\cdot, c) \text{ è } \varepsilon_c\text{-coerente}\}$$

dove $\varepsilon_c \geq 0$ è il parametro di coerenza nel contesto c . Una sezione globale $f \in \mathcal{F}^C(X)$ è un assegnamento che è ε_c -coerente simultaneamente in tutti i contesti.

Proposizione 47 (Tensione Contestuale come Coomologia del Fascio).

Il gruppo $H^1(X, \mathcal{F}^C)$ misura la tensione contestuale: l'impossibilità di estendere famiglie di assegnamenti localmente contestualmente compatibili a un assegnamento globale. $H^1(X, \mathcal{F}^C) = 0$ se e solo se ogni famiglia compatibile di assegnamenti locali si estende globalmente (la vaghezza contestuale non produce ostruzioni nuove rispetto alla vaghezza intra-contestuale).

Esempio. Il predicato 'è alto' con $C = \{\text{basket, ginnastica, quotidiano}\}$ ha soglie diverse per contesto: $\tau_{\{\text{basket}\}} \approx 192$ cm, $\tau_{\{\text{ginnastica}\}} \approx 170$ cm, $\tau_{\{\text{quotidiano}\}} \approx 180$ cm. Un parlante che si muove tra i contesti deve gestire assegnamenti localmente coerenti in ciascun contesto ma globalmente incompatibili. L'elemento non-zero di $H^1(X, \mathcal{F}^C)$ corrisponde a questa

incompatibilità globale: è la misura matematica precisa della confusione semantica prodotta dal cambio di contesto.

Capitolo 20. Dimostrazione Rigorosa del Lemma 1 e Struttura dei Cammini

Monotoni

Il Lemma 1 del Capitolo 5 afferma l'esistenza di un cammino monotono da $\min(X)$ a $\max(X)$ nel DAG $G^{\rightarrow}(X, \sim)$ quando $G(X, \sim)$ è connesso. La dimostrazione originale era schematica; questo capitolo la rende rigorosa con la tecnica delle inversioni, sviluppa risultati sulla struttura dello spazio dei cammini minimi, e presenta un algoritmo efficiente per la loro enumerazione.

20.1 Dimostrazione Rigorosa del Lemma 1

Lemma 1 (Riformulato con Ipotesi Minimali).

Sia $X = \{x_1 < x_2 < \dots < x_n\}$ un insieme totalmente ordinato e $G(X, \sim)$ il grafo di tolleranza con $\sim$ riflessiva e simmetrica. Se $G(X, \sim)$ è connesso, allora il DAG $G^{\rightarrow}(X, \sim)$ — ottenuto da $G(X, \sim)$ orientando ogni arco $\{x_i, x_j\}$ da x_i a x_j quando $i < j$ — contiene un cammino diretto da x_1 a x_n .

Lemma Ausiliare (Rettificazione). Per ogni coppia $x_i < x_j$ connessa da un cammino arbitrario π in $G(X, \sim)$, esiste un cammino monotono (non decrescente) da x_i a x_j in $G(X, \sim)$.

Dimostrazione del Lemma Ausiliare per induzione sul numero di inversioni $\text{inv}(\pi)$. Si chiama inversione di π ogni passo $z_k \rightarrow z_{k+1}$ con $z_{k+1} < z_k$.

Base ($\text{inv}(\pi) = 0$): π è già monotono. $\square$

Passo induttivo. Sia $\pi = (z_0, z_1, \dots, z_m)$ con $\text{inv}(\pi) = k \geq 1$.
Esiste

un indice j con $z_{j+1} < z_j$. Si distinguono tre casi:

Caso I: $z_{j-1} < z_{j+1}$ (il passo precedente a j è ascendente e può

'saltare' z_j). Si considera il cammino $\pi' = (z_0, \dots, z_{j-1}, z_{j+1}, \dots, z_m)$. Se $z_{j-1} \sim z_{j+1}$, π' è un cammino valido in $G(X, \sim)$ con $\text{inv}(\pi') \leq k-1$ (si è eliminata l'inversione j senza

aggiungerne). Se $z_{j-1} \not\sim z_{j+1}$, si nota che z_{j-1} e z_{j+1} sono connessi nel sottografo $G(X \cap [z_{j-1}, z_j], \sim)$ attraverso z_j ; per l'ipotesi induttiva su questo intervallo più piccolo, esiste un cammino monotono da z_{j-1} a z_{j+1} nel sottografo, che si usa per

sostituire z_{j-1}, z_j, z_{j+1} in π . Il cammino risultante π' ha $\text{inv}(\pi') \leq k-1$. Per ipotesi induttiva, da π' si ottiene un cammino

monotono.

Caso II: $z_{j-1} = z_{j+1}$. Allora si rimuove semplicemente il 'ciclo'

$z_{j-1}, z_j, z_{j+1} = z_{j-1}$ dal cammino, ottenendo $\pi' = (z_0, \dots, z_{j-1}, z_{j+2}, \dots, z_m)$ con $\text{inv}(\pi') \leq k-2$.

Caso III: $z_{j-1} > z_{j+1}$ (il 'picco' z_j è tra due valori entrambi più bassi). Si rimuove z_j se $z_{j-1} \sim z_{j+1}$; altrimenti si inserisce un cammino monotono da z_{j+1} a z_{j-1} (che esiste per ipotesi induttiva su $z_{j+1} < z_{j-1}$ connessi attraverso z_j)

invertendo il percorso. Il cammino risultante ha $\text{inv}(\pi') < k$.

In tutti i casi, il numero di inversioni decresce strettamente ad ogni

passo dell'induzione, e poiché $\text{inv}(\pi) \geq 0$ è intero, il processo

termina con $\text{inv} = 0$, cioè con un cammino monotono. Il Lemma 1 segue applicando il Lemma Ausiliare a $x_i = x_1, x_j = x_n$. $\square$

20.2 Struttura dello Spazio dei Cammini Minimi

Teorema 22 (Reticolo dei Cammini Minimi).

Lo spazio $\Pi_{\min}(X, \sim)$ dei cammini monotoni minimi da x_1 a x_n ha la struttura di un reticolo distributivo rispetto all'ordinamento puntuale $\pi \leq \pi'$ (l'i-esimo nodo di π è $\leq$ l'i-esimo nodo di π' per ogni i).

Precisamente:

(a) Per $\pi, \pi' \in \Pi_{\min}$, l'incontro $\pi \wedge \pi'$ (minimo puntuale) e l'unione $\pi \vee \pi'$ (massimo puntuale) sono ancora in $\Pi_{\min}$.

(b) Il reticolo è distributivo: $(\pi_1 \wedge \pi_2) \vee \pi_3 = (\pi_1 \vee \pi_3) \wedge (\pi_2 \vee \pi_3)$.

Dimostrazione. (a) Siano $\pi = (y_0, \dots, y_L)$ e $\pi' = (y'_0, \dots, y'_L)$ cammini minimi (stessa lunghezza $L = L_{\text{mono_min}}$). L'incontro puntuale è $z_j = \min(y_j, y'_j)$. Si verifica: (i) $z_0 = \min(y_0, y'_0) = \min(x_1, x_1) = x_1$ e $z_L = \min(y_L, y'_L) = x_n$. (ii) Monotonicità: $y_j \leq y_{j+1}$ e $y'_j \leq y'_{j+1}$ implicano $\min(y_j, y'_j) \leq \min(y_{j+1}, y'_{j+1})$. (iii) Adiacenza nel grafo: z_j e z_{j+1} sono entrambi nel cammino minimo da x_1 a x_n (il minimo puntuale di due cammini minimi è ancora un cammino minimo per un argomento di ottimalità nel DAG: se z non fosse un cammino minimo, esisterebbe un cammino più corto, contraddicendo la minimalità di π e π'). La stessa argomentazione vale per il massimo puntuale. (b) La distributività segue dalla struttura di reticolo di catena ($\mathbb{Z}$ con l'ordinamento naturale è distributivo, e il prodotto di reticoli distributivi è distributivo). $\square$

Proposizione 48 (Complessità dell'Enumerazione).

Il numero di cammini monotoni minimi può essere esponenziale in $L_{\text{mono_min}}$. Per la griglia $m \times n$, è il coefficiente binomiale $C(m+n-2, m-1)$, che cresce come $2^{\{m+n\}/\sqrt{m+n}}$ per $m, n \rightarrow \infty$. Per catene con tolleranza di larghezza k , il numero di cammini minimi è al massimo $k^{L_{\text{mono_min}}}$. Dunque l'enumerazione completa è intrinsecamente esponenziale, ma il calcolo del singolo cammino minimo rimane polinomiale in $O(|X| + |E|)$ tramite Dijkstra sul DAG.

20.3 Algoritmo di Enumerazione dei Cammini Minimi

Si estende l'algoritmo di programmazione dinamica per il calcolo di $L_{\text{mono_min}}$ (Proposizione 12, Capitolo 5) alla produzione di tutti i cammini minimi.

Algoritmo EnumeraCamminiMinimi($X, \sim$):

1. Calcola $d[x_i] = L_mono_min$ da x_1 a x_i per ogni i con visita BFS sul DAG ($O(|X| + |E|)$).
2. Per ogni x_i , calcola $Pred[x_i] = \{x_j : x_j \sim x_i, j < i, d[x_j] = d[x_i] - 1\}$ (l'insieme dei predecessori ottimali di x_i nel DAG dei cammini minimi).
3. Enumera tutti i cammini da x_n a x_1 percorrendo il grafo di predecessori con DFS a ritroso, invertendo l'ordine dei nodi. Ogni percorso DFS produce un cammino minimo distinto.

Complessità: $O(|X| + |E| + |output|)$, dove $|output|$ è il numero totale di nodi in tutti i cammini minimi enumerati. L'algoritmo è output-sensitive: se ci sono K cammini minimi di lunghezza L , la complessità è $O(|X| + |E| + K \cdot L)$.

Capitolo 21. Teoria di Morse Discreta per Complessi di Sorite

La teoria di Morse discreta, sviluppata da Robin Forman (1998, 2002), è il corrispettivo combinatorio della teoria di Morse per varietà differenziabili. Fornisce strumenti per semplificare un complesso simpliciale riducendolo a un CW-complesso di dimensione minima con lo stesso tipo di omotopia, identificando i "simplessi critici" che generano l'omologia. Applicata al complesso di Sorite, la teoria di Morse discreta localizza topologicamente la struttura di vaghezza e fornisce nuovi algoritmi computazionali per il sorites barcode.

21.1 Funzioni di Morse discrete e campi vettoriali

Definizione 21.1 (Funzione di Morse discreta). Una funzione di Morse discreta su un complesso simpliciale K è una mappa $f: K \rightarrow \mathbb{R}$ che assegna un valore reale a ogni semplice, soddisfacendo: per ogni k -simpleso $\sigma \in K$, (i) il numero di facce $(k-1)$ -dimensionali $\tau < \sigma$ con $f(\tau) \geq f(\sigma)$ è al più 1; (ii) il numero di cofacce $(k+1)$ -dimensionali $\upsilon > \sigma$ con $f(\upsilon) \leq f(\sigma)$ è al più 1. Un semplice σ è critico se nessuna delle due condizioni si verifica: $f(\tau) < f(\sigma)$ per tutte le facce $\tau < \sigma$, e $f(\upsilon) > f(\sigma)$ per tutte le cofacce $\upsilon > \sigma$.

Il teorema di Forman afferma che il numero di simplessi critici di dimensione k è almeno $\beta_k(K)$, il k -esimo numero di Betti. Un complesso con esattamente β_k simplessi critici di dimensione k per ogni k è detto complesso di Morse perfetto. Per i complessi di Sorite, la costruzione di una funzione di Morse ottimale è direttamente collegata alla struttura del cammino monotono minimo.

Teorema 21.1 (Funzione di Morse canonica sul complesso di Sorite). Sia $(X, \sim, \leq)$ uno spazio di tolleranza totalmente ordinato con $G(X, \sim)$ connesso. La funzione $f^*: S(X, \sim) \rightarrow \mathbb{R}$

definita su ogni simpleso $\sigma = \{x_{\{i_0\}}, x_{\{i_1\}}, \dots, x_{\{i_k\}}\}$ (con $i_0 < i_1 < \dots < i_k$) da $f^*(\sigma) = x_{\{i_0\}} + k \cdot \varepsilon^*$ è una funzione di Morse discreta sul complesso di Sorite $S(X, \sim)$. I semplici critici sono: (i) il 0-simpleso $\{\min(X)\}$ (minimo globale); (ii) il 0-simpleso $\{\max(X)\}$ (massimo relativo, generatore di H_0); (iii) gli 1-simplessi $\{y_{\{j-1\}}, y_{\{j\}}\}$ per $j = 1, \dots, L_{\text{mono_min}}$ che costituiscono il cammino monotono minimo (generatori di cicli se il complesso ha $H_1 \neq 0$ per struttura ciclica).

Dimostrazione. Si verifica la condizione di Morse per f^* su ogni simpleso. Per un 0-simpleso $\{x\}$: le sue cofacce $\{x, y\}$ per $x < y$ soddisfano $f^*(\{x, y\}) = x + \varepsilon^* > x = f^*(\{x\})$ solo se $\varepsilon^* > 0$, il che vale. Il 0-simpleso $\{\min(X)\}$ non ha facce e ha tutte le cofacce con valore maggiore: è critico. Per un 1-simpleso $\{x_i, x_j\}$ con $i < j$: la faccia $\{x_i\}$ ha $f^* = x_i < x_i + \varepsilon^* = f^*(\{x_i, x_j\})$; la faccia $\{x_j\}$ ha $f^* = x_j > x_i + \varepsilon^*$ quando $x_j - x_i > \varepsilon^*$. Le cofacce del 1-simpleso sono i 2-simplessi che lo contengono. Poiché il complesso di Sorite unidimensionale (grafi senza triangoli) non ha 2-simplessi, tutti gli 1-simplessi sono critici o accoppiati a 0-simplessi. Il pairing è dato dal gradiente discreto. $\square$

21.2 Il campo vettoriale di Morse e il complesso di Morse

Definizione 21.2 (Campo vettoriale discreto). Un campo vettoriale discreto V su un complesso simpliciale K è una collezione di coppie (σ, τ) dove σ è un $(k-1)$ -simpleso, τ è un k -simpleso, $\sigma < \tau$ (σ è faccia di τ), e ogni simpleso compare in al più una coppia. Le V -cammini sono le sequenze $\sigma_0, \tau_0, \sigma_1, \tau_1, \dots, \sigma_m$ di alternanze faccia-cofaccia dove $(\sigma_i, \tau_i) \in V$ e $\sigma_{i+1} < \tau_i$, $\sigma_{i+1} \neq \sigma_i$. Un campo vettoriale è senza V -orbite chiuse.

Il teorema fondamentale di Forman (1998) stabilisce che un campo vettoriale discreto senza V -orbite chiuse corrisponde biunivocamente a una funzione di Morse discreta (con gli accoppiamenti dati dalle coppie del campo), e che il tipo di omotopia di K coincide con quello del CW-complesso ottenuto da un punto aggiungendo una cella di dimensione k per ogni simpleso critico di dimensione k .

Per il complesso di Sorite di un predicato unidimensionale con $G(X, \sim)$ connesso, il campo vettoriale di Morse canonico è costruito come segue: si accoppia ogni 0-simpleso $\{x_i\}$ con l'arco $\{x_{i-1}, x_i\}$ lungo il cammino monotono minimo, lasciando i soli semplici critici $\{x_1\}$ e ogni arco fuori dal cammino minimo privo di accoppiamento. Il numero di semplici critici è: 1 (il minimo $\{x_1\}$) più $|E_{\text{out}}|$ (gli archi fuori dal cammino minimo). La relazione con i numeri di Betti: $\beta_0 = 1$ e $\beta_1 = |E_{\text{out}}| - (n-1-L_{\text{monomin}})$ per la struttura del grafo.

21.3 Le disuguaglianze di Morse per il complesso di Sorite

Teorema 21.2 (Disuguaglianze di Morse per spazi di tolleranza). Sia $K = S(X, \sim)$ il complesso di Sorite. Denotiamo con c_k il numero di semplici critici di dimensione k di una qualunque funzione di Morse discreta su K . Valgono le disuguaglianze di Morse:

$$c_k \geq \beta_k(K) \text{ per ogni } k \geq 0$$

$$\sum_k (-1)^k c_k = \sum_k (-1)^k \beta_k = \chi(K) \text{ (caratteristica di Euler)}$$

In particolare: $c_0 \geq \beta_0 = 1$ (almeno un semplice critico di dimensione 0, il minimo globale). Per grafi connessi con μ cicli indipendenti: $c_1 \geq \beta_1 = \mu$ (almeno μ archi critici). La differenza $c_k - \beta_k$ è pari se non nulla (si annulla a coppie), per il teorema di Morse forte.

Dimostrazione. Il risultato è il Teorema 2.4 di Forman (1998). La dimostrazione usa la struttura della catena di Morse: i semplici critici generano un complesso di catene di Morse C^*_Morse con differenziale di Morse che conta i V-cammini tra semplici critici. Il complesso di Morse ha la stessa omologia di K per il teorema di omotopia, e la formula degli indici di Morse si ottiene applicando il lemma del serpente alla sequenza delle filtrazioni. $\square$

21.4 Applicazione: semplificazione del sorites barcode tramite collasso elementare

La teoria di Morse discreta fornisce algoritmi di semplificazione del complesso di Sorite basati sul concetto di collasso elementare: una coppia (σ, τ) con σ faccia libera di τ (τ è l'unica cofaccia di σ nel complesso) può essere collassata, eliminando entrambi i semplici senza cambiare il tipo di omotopia. Il processo di collasso iterativo è equivalente all'applicazione del campo vettoriale di Morse.

Proposizione 21.1 (Collasso del sorites complex per predicati unidimensionali). Per il complesso di Sorite di un predicato unidimensionale $S(X, \sim)$ con $G(X, \sim)$ connesso e senza cicli (un albero di copertura), il complesso si collassa a un punto tramite L_mono_min collassi elementari sequenziali. Il sorites barcode PD_0 ha esattamente $n-1$ barre, una per ogni collasso, e il sorites threshold r^* corrisponde all'ultimo collasso (la barra di massima persistenza).

Questo risultato ha un'importante conseguenza algoritmica: il sorites barcode per predicati unidimensionali può essere calcolato in tempo $O(n \log n)$ tramite un algoritmo di spanning tree minimo (Kruskal o Prim), senza la necessità di calcolare l'intera omologia persistente. Il costo computazionale è molto inferiore all'algoritmo generale $O(m^3)$ per complessi simpliciali.

21.5 Morse theory e il problema della localizzazione delle zone di confine

La teoria di Morse discreta fornisce una risposta precisa alla domanda semantica "dove si trova il confine del predicato?". I semplici critici di dimensione 1 (gli archi critici, non accoppiati dal campo vettoriale) sono esattamente i "passi critici" della classificazione: gli archi del grafo di tolleranza che, se rimossi, aumenterebbero il numero di componenti connesse e quindi permetterebbero classificazioni coerenti più fini.

In particolare, per il predicato "è alto" su una catena lineare, gli archi critici si concentrano nella zona intorno alla media della distribuzione delle stature, dove la densità di relazioni di tolleranza è più alta. Per predicati multidimensionali come "è in forma", gli archi critici si distribuiscono su ipersuperfici nel prodotto cartesiano degli spazi unidimensionali, producendo "iperplani di confine" multipli corrispondenti alle $(m-1)(n-1)$ ostruzioni omologiche della griglia.

Capitolo 22. Coomologia con Prodotti Cup e Teorema di Alexander

La coomologia simpliciale, introdotta nel Capitolo 10, ha una struttura algebrica più ricca dell'omologia: oltre alla struttura di gruppo abeliano, porta un prodotto bilineare chiamato prodotto cup, che conferisce all'algebra coomologica la struttura di un'algebra graduata. Questo prodotto cattura informazioni topologiche che l'omologia da sola non può vedere: in particolare, la "lunghezza cup" (il massimo numero di fattori non banali in un prodotto cup non nullo) è un invariante superiore strettamente connesso alla complessità della struttura di vaghezza. Questo capitolo sviluppa la teoria del prodotto cup per i complessi di Sorite e ne trae conseguenze semantiche.

22.1 Il prodotto cup in coomologia simpliciale

Definizione 22.1 (Prodotto cup). Per un complesso simpliciale K con orientamento fisso dei semplici, il prodotto cup $\smile: C^p(K;R) \times C^q(K;R) \rightarrow C^{p+q}(K;R)$ è definito su cocatene $\varphi \in C^p$ e $\psi \in C^q$ da: $(\varphi \smile \psi)(\sigma) = \varphi([v_0, \dots, v_p]) \cdot \psi([v_p, \dots, v_{p+q}])$ per ogni $(p+q)$ -simpleso $\sigma = [v_0, \dots, v_{p+q}]$ (con vertici ordinati). Il prodotto cup è associativo e soddisfa la formula di Leibniz rispetto al cobordo: $\delta(\varphi \smile \psi) = (\delta\varphi) \smile \psi + (-1)^p \varphi \smile (\delta\psi)$. Dunque il prodotto cup discende a un prodotto well-defined in coomologia: $H^*(K;R) \times H^*(K;R) \rightarrow H^*(K;R)$.

Il prodotto cup è graduato-commutativo: per $\alpha \in H^p$ e $\beta \in H^q$, vale $\alpha \smile \beta = (-1)^{pq} \beta \smile \alpha$. La struttura di anello su $H^*(K;R)$ ottenuta dal prodotto cup è un invariante topologico del tipo di omotopia di $|K|$ più fine dell'omologia dei singoli gradi.

22.2 Il prodotto cup per complessi di Sorite

Proposizione 22.1 (Prodotto cup per i complessi di flag). Per un complesso di flag $K = S(X, \sim)$ (complesso delle clique di $G(X, \sim)$), il prodotto cup $H^1(K;R) \times H^1(K;R) \rightarrow H^2(K;R)$ è non banale se e solo se K ha $H_2 \neq 0$, ovvero se e solo se il grafo di tolleranza $G(X, \sim)$ contiene una struttura triangolare con "vuoto 2-dimensionale" (un ciclo di 2-simplessi che non è bordo di un 3-simpleso).

Dimostrazione. Per un complesso 1-dimensionale (grafi senza triangoli), $H_2 = 0$ e il prodotto cup è banale. Per complessi 2-dimensionali, si usa la formula del prodotto cup su cochains: $(\varphi \smile \psi)(v_0, v_1, v_2) = \varphi(v_0, v_1) \cdot \psi(v_1, v_2)$. Se φ e ψ sono 1-cocicli che generano H^1 , il loro prodotto cup è non nullo in H^2 se e solo se esiste un triangolo (2-simpleso) $\{v_0, v_1, v_2\}$ su cui entrambe le restrizioni siano non-zero simultaneamente. Per complessi di flag con

struttura di clique appropriata, questo accade esattamente quando il grafo contiene strutture toroidali o sferiche non riempibili. $\square$

Per i complessi di Sorite dei predicati linguistici naturali, la Congettura 4 del Capitolo 14 afferma $H_2 = 0$. Se corretta, il prodotto cup $H^1 \times H^1 \rightarrow H_2$ è identicamente zero, e l'algebra coomologica è completamente determinata da H^0 , H^1 con prodotto cup banale in grado 2. Questo è coerente con la struttura "planare" della vaghezza percettiva.

22.3 La cup-length come invariante della struttura di vaghezza

Definizione 22.2 (Cup-length). La cup-length di un complesso K su un anello R è $cl(K; R) = \max \{k : \exists \alpha_1, \dots, \alpha_k \in H^*(K; R) \text{ con } |\alpha_i| \geq 1 \text{ per ogni } i \text{ e } \alpha_1 \smile \dots \smile \alpha_k \neq 0\}$. La cup-length è un invariante del tipo di omotopia di $|K|$, più fine dei numeri di Betti.

Proposizione 22.2 (Cup-length e complessità della vaghezza). Per il complesso di Sorite $S(X, \sim)$ di un predicato multidimensionale su $X^d = X_1 \times \dots \times X_d$ con tolleranza prodotto e con complesso di Sorite omotopicamente equivalente al toro $T^d = S^1 \times \dots \times S^1$ (d fattori):

$$cl(S(X, \sim); R) = d$$

La cup-length d corrisponde alla dimensione concettuale del predicato: predicati genuinamente d -dimensionali hanno cup-length d , e nessun insieme di $d-1$ generatori di H^1 ha prodotto cup non nullo. Questo fornisce un invariante algebrico preciso della dimensione semantica.

22.4 Il teorema di dualità di Alexander e le ostruzioni globali

Teorema 22.1 (Dualità di Alexander per complessi di Sorite). Sia $K = S(X, \sim)$ un complesso simpliciale contenuto nell' $(n-1)$ -simpleso standard Δ^{n-1} (dove $n = |X|$). Il complemento di Alexander è $K^c = \Delta^{n-1} \setminus |K|$. La dualità di Alexander stabilisce:

$$H_k^{\sim}(K; \mathbb{Z}) \cong H_{n-k-2}^{\sim}(K^c; \mathbb{Z}) \text{ per ogni } k$$

dove $\tilde{H}$ denota l'omologia (risp. coomologia) ridotta. Per i complessi di Sorite, il complemento di Alexander K^c corrisponde ai sottoinsiemi "non-tolleranti" di X : quelli che non sono clique del grafo $G(X, \sim)$. La dualità di Alexander stabilisce una corrispondenza biunivoca tra le ostruzioni omologiche k -dimensionali di $S(X, \sim)$ e le ostruzioni coomologiche $(n-k-2)$ -dimensionali della struttura dei "non-tolleranti".

Dimostrazione. Il Teorema di Dualità di Alexander è Teorema 71.1 di Munkres (1984). La sua applicazione al contesto degli spazi di tolleranza richiede l'embedding naturale di $S(X, \sim)$ in Δ^{n-1} dato dall'inclusione del complesso simpliciale nel simpleso pieno sull'insieme X . Il complemento K^c è il sottocomplesso di Δ^{n-1} formato dai semplici che non sono clique di $G(X, \sim)$: $\{\sigma \subseteq X : \sigma \text{ non è mutualmente tollerante}\} = \{\sigma : \exists x_i, x_j \in \sigma \text{ con } x_i \not\sim x_j\}$. $\square$

Interpretazione semantica della dualità di Alexander. Le ostruzioni omologiche di primo ordine $H_1(S(X, \sim)) = \mathbb{Z}^m$ (i cicli del grafo di tolleranza) sono dual alle ostruzioni coomologiche di ordine $n-3$ del "complemento di non-tolleranza". Per predicati unidimensionali (grafi planari), $n-3 = n-3$ è grande, e le ostruzioni del complemento sono in omologia alta: l'omologia alta del complemento misura la struttura globale degli elementi "isolati" (quelli con pochi vicini tolleranti). Questo fornisce una dualità tra la struttura locale del paradosso (gli elementi al confine con molti vicini tolleranti) e la struttura globale degli elementi paradigmatici (con pochi vicini).

22.5 Calcolo esplicito: la dualità di Alexander per C_4

Si calcola la dualità di Alexander per $K = S(C_4, \sim_1) \cong S^1$ con $X = \{1, 2, 3, 4\}$, $n = 4$. Il simpleso standard è Δ^3 (il tetraedro). Il complemento $K^c \subseteq \Delta^3$ contiene i semplici che non sono clique di C_4 : le coppie $\{1, 3\}$ e $\{2, 4\}$ (le diagonali, non archi in C_4), e tutti i semplici che li contengono, cioè le triple $\{1, 3, 2\}$, $\{1, 3, 4\}$ e le quadruple $\{1, 2, 3, 4\}$.

Il complemento K^c ha 4 vertici, 1 arco $\{1, 3\}$, 1 arco $\{2, 4\}$, e i 4 triangoli che contengono tali archi. L'omologia di K^c : le due diagonali $\{1, 3\}$ e $\{2, 4\}$ si incontrano in? Nel Δ^3 tridimensionale, $\{1, 3\}$ e $\{2, 4\}$ non condividono vertici, dunque K^c è il join di due archi disassociati. Il tipo di omotopia di K^c è $S^0 \times [0, 1] / \sim = S^1$. La dualità di Alexander verifica: $\tilde{H}_1(S(C_4, \sim_1)) = \mathbb{Z} \cong \tilde{H}^0(K^c) = \mathbb{Z}$. Coerente: $\tilde{H}^0(K^c) = \mathbb{Z}$ perché K^c ha 2 componenti connesse (i due archi disassociati), dunque $\beta_0(K^c) = 2$ e $\tilde{H}^0(K^c) = \mathbb{Z}$.

Capitolo 23. Il Mapper Algorithm per la Visualizzazione della Vaghezza

Il Mapper è un algoritmo di topological data analysis sviluppato da Singh, Mémoli e Carlsson (2007) che produce una rappresentazione grafica compatta della "forma" di un insieme di dati ad alta dimensione. Applicato agli spazi di tolleranza, il Mapper produce una visualizzazione topologica della struttura di vaghezza di un predicato che è più interpretabile del sorites barcode

pur essendo meno formale. Questo capitolo sviluppa la versione del Mapper per gli spazi di tolleranza e ne discute le proprietà teoriche.

23.1 La costruzione di Mapper

Definizione 23.1 (Costruzione di Mapper). Siano $(X, \sim)$ uno spazio di tolleranza finito, $f: X \rightarrow \mathbb{R}$ una "funzione filtro" (tipicamente $f(x)$ = valore del predicato o distanza da un punto di riferimento), e $U = \{U_a\}_a$ un ricoprimento aperto dell'immagine $f(X) \subseteq \mathbb{R}$ da intervalli aperti sovrapposti. Il Mapper $M(X, f, U)$ è il grafo costruito come segue: (i) i vertici di M sono le componenti connesse (rispetto alla tolleranza $\sim$) degli insiemi $f^{-1}(U_a)$ per ogni a ; (ii) due vertici sono collegati da un arco se i corrispondenti sottoinsiemi di X si sovrappongono.

Il Mapper è una versione discreta del teorema del nervo: il grafo M approssima il tipo di omotopia di X rispetto alla funzione filtro f , con una risoluzione che dipende dalla scelta del ricoprimento U . Ricoprimenti più fini producono un Mapper più dettagliato; ricoprimenti più grossolani danno una visione più aggregata.

23.2 Il Mapper per la vaghezza

Per il predicato "è alto" su $X = \{\text{stature campione}\}$ con $f(x) = x$ (la statura in cm) e ricoprimento U dato da intervalli $[a, a+w]$ sovrapposti di lunghezza w e overlap δ , il Mapper produce un grafo che visualizza la struttura di clustering delle stature sotto la relazione di tolleranza. Tipicamente, il Mapper mostra: (i) una singola struttura lineare per predicati unidimensionali semplici (catena di cluster di stature simili); (ii) strutture ramificate per predicati con zone di transizione multiple; (iii) strutture anulari per predicati ciclici.

Proposizione 23.1 (Convergenza del Mapper al sorites barcode). Per ricoprimenti U_ϵ con intervalli di lunghezza $w \rightarrow 0$ e overlap $\delta \rightarrow 0$, il grafo Mapper $M(X, f, U_\epsilon)$ converge (in senso di Gromov-Hausdorff) alla realizzazione geometrica del 1-scheletro di $S(X, \sim_w)$, il complesso di Sorite al raggio di tolleranza $r = w$. In questo limite, le componenti connesse del Mapper coincidono con le componenti connesse di $G(X, \sim_w)$, e il sorites threshold r^* può essere letto direttamente dalla struttura del Mapper per ricoprimenti appropriati.

Dimostrazione. La dimostrazione usa la costruzione del nervo: $M(X, f, U_\epsilon)$ è il nervo del ricoprimento $\{f^{-1}(U_a)\}_a$ del grafo di tolleranza $G(X, \sim_w)$. Per il Teorema del Nervo (Capitolo 3), questo nervo ha lo stesso tipo di omotopia di $G(X, \sim_w)$ se ogni intersezione finita degli elementi del ricoprimento è contraibile (o connessa). Per ricoprimenti sufficientemente fini, ogni intersezione $f^{-1}(U_a) \cap f^{-1}(U_b)$ è un sottoinsieme di X contenuto in un intervallo di stature di larghezza $2w$, che è contraibile per la struttura lineare di X . $\square$

23.3 Algoritmo e complessità del Mapper per predicati vaghi

L'algoritmo del Mapper per spazi di tolleranza è il seguente: (1) Calcola $f: X \rightarrow \mathbb{R}$ (la funzione filtro, tipicamente la posizione ordinale o una proiezione PCA per predicati multidimensionali). (2) Costruisci il ricoprimento $U = \{U_a\}$ dell'immagine $f(X)$: scegli n_{interval} intervalli sovrapposti con overlap del $\text{poverlap}\%$ (tipicamente 10–30%). (3) Per ogni U_a , estrai il sottoinsieme $X_a = \{x \in X : f(x) \in U_a\}$ e calcola le sue componenti connesse sotto la relazione $\sim$. (4) Crea i vertici del Mapper (uno per componente) e gli archi (tra componenti che si sovrappongono).

La complessità è $O(n_{\text{interval}} \cdot n^2)$ nel caso peggiore ($n = |X|$), ma $O(n_{\text{interval}} \cdot n \cdot d_{\text{max}})$ dove d_{max} è il grado massimo del grafo di tolleranza per grafi sparsi. In pratica, per predicati unidimensionali con tolleranza di larghezza δ , il grafo è sparso con $d_{\text{max}} = 2\delta/\Delta x$ dove Δx è la spaziatura degli elementi, e il Mapper è calcolabile in tempo quasi-lineare.

23.4 Esempio: il Mapper per il predicato "è calvo"

Si applica il Mapper al predicato "è calvo" su $X = \{0, 1000, 2000, \dots, 100000\}$ (101 valori del numero di capelli) con tolleranza $\sim_\delta$ di raggio $\delta = 10000$ capelli. La funzione filtro è $f(x) = x$ (il numero di capelli). Il ricoprimento U consiste di 10 intervalli di larghezza 15000 con overlap del 30%: $U_1 = [-2250, 12750]$, $U_2 = [7750, 27750]$, ..., $U_{10} = [87750, 107750]$.

Il Mapper risultante è un grafo lineare di 10 vertici (uno per ogni intervallo, con una sola componente connessa per la struttura lineare della catena con $\delta = 10000$). Ogni arco del Mapper corrisponde alla transizione tra cluster di stature adiacenti. Il sorites threshold si legge come il numero di archi del Mapper ($10 - 1 = 9$ transizioni), e $\varepsilon^* = 1/L_{\text{monomin}} \approx 1/10 = 0,1$ (approssimativamente, dipende dalla granularità del ricoprimento). Per predicati bidimensionali, il Mapper produce un grafo con struttura "a rete" che riflette le $(m-1)(n-1)$ ostruzioni omologiche della griglia.

Capitolo 24. Moduli Persistenti su Quiver e Classificazione dei Predicati Vaghi

I moduli persistenti generalizzano i diagrammi di persistenza sostituendo la struttura lineare dei raggi r con strutture di quiver più generali. Questa generalizzazione è rilevante per i predicati vaghi multidimensionali: la filtrazione di Sorite su una griglia non è indicizzata da un singolo parametro reale ma da un reticolo bidimensionale di parametri (r_i, r_j) corrispondenti alle soglie di tolleranza nelle diverse dimensioni. Il teorema di struttura dei moduli persistenti su quiver fornisce condizioni di classificabilità e invarianti completi.

24.1 Moduli persistenti su poset

Definizione 24.1 (Modulo persistente su poset). Sia $(P, \leq)$ un poset. Un modulo persistente su P con valori in VettR (la categoria dei spazi vettoriali reali) è un funtore $M: P \rightarrow \text{VettR}$. Esplicitamente: per ogni $p \in P$, $M(p)$ è uno spazio vettoriale; per ogni $p \leq q$, $M(p \leq q): M(p) \rightarrow M(q)$ è un'applicazione lineare (la mappa di struttura); le mappe di struttura sono compatibili con la composizione: $M(q \leq r) \circ M(p \leq q) = M(p \leq r)$.

Per la filtrazione di Sorite unidimensionale, $P = (\mathbb{R}, \leq)$ con le mappe di struttura date dall'inclusione $S(X, \sim_r) \subseteq S(X, \sim_s)$ per $r \leq s$. Il modulo persistente associato è $M_k(r) = H_k(S(X, \sim_r); \mathbb{R})$. Per la filtrazione di Sorite bidimensionale su griglia, $P = (\mathbb{R} \times \mathbb{R}, \leq_{\text{componentwise}})$ con $M_k(r_1, r_2) = H_k(S(X, \sim_{\{r_1, r_2\}}); \mathbb{R})$, dove $\sim_{\{r_1, r_2\}}$ è la tolleranza del prodotto con raggi (r_1, r_2) nelle due dimensioni.

24.2 Il teorema di struttura per il caso lineare

Teorema 24.1 (Teorema di struttura per moduli persistenti su $\mathbb{R}$, Crawley-Boevey 2012). Ogni modulo persistente M su $\mathbb{R}$ con valori in VettR che soddisfa la condizione q -tame ($\dim M(r) < \infty$ per ogni $r \in \mathbb{R}$, e le mappe di struttura sono isomorfismi fuori da un insieme finito di valori) è isomorfo come modulo persistente a una somma diretta di moduli intervallo:

$$M \cong \bigoplus_i R[a_i, b_i)$$

dove $R[a, b)$ è il modulo intervallo che vale $\mathbb{R}$ per $a \leq r < b$ e 0 altrimenti. L'insieme $\{[a_i, b_i)\}$ è il diagramma di persistenza del modulo M , ed è un invariante completo del modulo (a isomorfismo). Questo teorema giustifica il formalismo del diagramma di persistenza: ogni classe omologica persistente è descritta esattamente da un intervallo di nascita/morte.

24.3 L'ostacolo alla classificazione in dimensione 2: il teorema di

Gabriel

Teorema 24.2 (Teorema di Gabriel, versione TDA, Carlsson-Zomorodian 2009). Il problema di classificazione dei moduli persistenti biparametrici (funtori da $\mathbb{R}^2$ a VettR) non ammette una decomposizione finita analoga a quella del caso lineare: esistono moduli biparametrici indecomponibili di qualsiasi dimensione, e il problema di classificazione è di

tipo "tipo selvatico" (wild type) nel linguaggio della teoria delle rappresentazioni dei quiver.

Il Teorema 24.2 ha implicazioni dirette per la vaghezza dei predicati multidimensionali: non esiste un invariante numerico finito che classifichi completamente la struttura di vaghezza di un predicato bivariato modulo equivalenza isometrica. Il sorites barcode (la somma dei diagrammi di persistenza $PD_0, PD_1, \dots$) è incompleto per predicati genuinamente multidimensionali: due predicati bivariati con lo stesso barcode in tutti i gradi possono avere strutture di vaghezza topologicamente diverse.

Questo risultato formalizza matematicamente l'intuizione semantica che i predicati multidimensionali ("è bello", "è intelligente", "è in forma") abbiano una struttura di vaghezza irriducibile alle componenti unidimensionali: la topologia bidimensionale introduce una complessità qualitativa che non è catturabile da invarianti unidimensionali.

24.4 Invarianti per moduli persistenti biparametrici: la funzione di

Hilbert e i numeri di Betti bigraduati

Benché la classificazione completa sia impossibile, esistono invarianti parziali utili per i moduli persistenti biparametrici. I più importanti sono:

La funzione di Hilbert: $HM(r_1, r_2) = \dim M(r_1, r_2)$. Per il complesso di Sorite bidimensionale, $H\{M_i\}(r_1, r_2)$ è il profilo di Betti bidimensionale $\beta_1(r_1, r_2) = \text{rk}(H_1(S(X, \sim_{\{r_1, r_2\}})))$.

I numeri di Betti bigraduati: Per ogni coppia (r_1, r_2) con $r_1 \leq r_2$, il numero di Betti bigraduato è $\xi_{ij} = \dim(\ker(M_i \rightarrow M_{i+1}) \cap \ker(M_i \rightarrow M_j)) / \dim(M_{i-1} \rightarrow M_i) + \dim(M_{j+1} \rightarrow M_j)$. I numeri di Betti bigraduati sono invarianti dei moduli biparametrici che generalizzano i punti del diagramma di persistenza unidimensionale.

Proposizione 24.1 (Numeri di Betti bigraduati per la griglia). Per il complesso di Sorite della griglia $m \times n$ con filtrazione biparametrica (r_1, r_2) corrispondente alle soglie nelle due dimensioni, il numero di Betti bigraduato $\xi_{ij}(k)$ in grado $k = 0$ è:

$$\xi_{ij}(0) = \beta_0(r_{1i}, r_{2j}) - \beta_0(r_{1, i+1}, r_{2j}) - \beta_0(r_{1i}, r_{2, j+1}) + \beta_0(r_{1, i+1}, r_{2, j+1})$$

dove $\beta_0(r_1, r_2)$ è il numero di componenti connesse del grafo di tolleranza con soglia (r_1, r_2) . Questa formula è la versione discreta della derivata mista della funzione di Hilbert, e conteggia le nascite "genuine" di componenti connesse al raggio (r_{1i}, r_{2j}) , al netto di quelle ereditate dai raggi inferiori.

Capitolo 25. Estensioni Metriche: La Distanza di Interleaving tra predicati

Il teorema di stabilità del sorites barcode (Teorema 9.1) misura la stabilità in termini della distanza bottleneck d_B . Esiste però una nozione alternativa di distanza tra moduli persistenti — la distanza di interleaving — che è più naturale dal punto di vista categoriale e ha proprietà teoriche più forti.

Definizione 25.1 (Interleaving di moduli persistenti). Due moduli persistenti M e N su $\mathbb{R}$ sono δ -interlacciati (δ -interleaved) se esistono morfismi di moduli $\varphi: M \rightarrow N(\delta)$ e $\psi: N \rightarrow M(\delta)$ (dove $M(\delta)$ è il modulo traslato: $M(\delta)(r) = M(r+\delta)$) tali che $\psi(\delta) \circ \varphi = M(r \leq r+2\delta)$ e $\varphi(\delta) \circ \psi = N(r \leq r+2\delta)$. La distanza di interleaving è $d_I(M, N) = \inf\{\delta \geq 0 : M \text{ e } N \text{ sono } \delta\text{-interlacciati}\}$.

Teorema 25.1 (Isometry Theorem, Lesnick 2015). Per moduli persistenti q-tame su $\mathbb{R}$, la distanza di interleaving e la distanza bottleneck tra i diagrammi di persistenza coincidono:

$$d_I(M, N) = d_B(PD(M), PD(N))$$

L'Isometry Theorem è uno dei risultati più profondi della TDA: mostra che la distanza bottleneck, definita combinatorialmente tra diagrammi di persistenza, ha un'interpretazione categoriale precisa come distanza tra funtori. Per i complessi di Sorite, questo significa che due predicati con distanza bottleneck ε tra i loro barcode sono ε -interlacciati: esiste una "traduzione" tra le strutture di tolleranza che sposta i raggi di al più ε .

Capitolo 26. Statistica Inferenziale del Sorites Barcode: Bootstrap, Convergenza e Test di Ipotesi

26.1 Il problema dell'inferenza topologica

Quando si estrae il sorites barcode da dati empirici — siano essi giudizi di classificazione raccolti via survey, distanze coseno calcolate da embedding neurali, o misure psicofisiche di indistinguibilità — ci si trova inevitabilmente di fronte a un campione finito di dimensione n estratto da una popolazione virtualmente infinita di parlanti o di esperienze percettive. Sorge allora una questione fondamentale che i capitoli precedenti non hanno affrontato: quanto può essere distante il barcode calcolato sul campione da quello che si otterrebbe conoscendo l'intera struttura di tolleranza della comunità linguistica? In altri termini, quale è l'incertezza statistica associata al sorites barcode come stimatore di una quantità semantica reale?

Questo capitolo affronta il problema dell'inferenza statistica per diagrammi di persistenza e sorites barcode, sviluppando strumenti che permettono di costruire intervalli di confidenza, testare ipotesi sulla struttura topologica della vaghezza, e quantificare la tasso di convergenza del barcode campionario al barcode della distribuzione sottostante. I risultati di questo capitolo sono fondamentali per rendere la teoria degli spazi di tolleranza empiricamente verificabile: senza stime di incertezza, il sorites barcode è una descrizione del campione specifico analizzato, non un'inferenza sulla struttura semantica del predicato.

La struttura matematica dell'inferenza topologica combina risultati della statistica non parametrica (bootstrap, test di permutazione), della geometria stocastica (convergenza delle filtrazioni di Čech e Vietoris-Rips al crescere del campione), e della teoria della misurazione dei diagrammi di persistenza (spazi di Wasserstein, metriche bottleneck). Il punto di partenza è il risultato di Chazal, Fasy, Lecci, Rinaldo, Singh e Wasserman (2014), che fornisce il quadro teorico per la costruzione di bande di confidenza per i diagrammi di persistenza.

Formalizziamo il problema. Si fissi un predicato P con struttura di tolleranza genuina su una popolazione sottostante. Sia μ una misura di probabilità su X che descrive la distribuzione delle istanze del predicato nella popolazione. Il sorites barcode "vero" del predicato è $PD^*(k) = PD_k(S(X, \sim_r))_{\{r \geq 0\}}$, dove la filtrazione è quella indotta dalla struttura di tolleranza metrica della popolazione intera. In pratica, si osserva un campione finito $X_n = \{x_1, \dots, x_n\}$ i.i.d. da μ , e si calcola il sorites barcode empirico $PD_n(k) = PD_k(S(X_n, \sim_r))_{\{r \geq 0\}}$. La domanda statistica è: quanto è distante $PD_n(k)$ da $PD^*(k)$?

Definizione 26.1 (Spazio di Wasserstein per diagrammi di persistenza). Lo spazio W_p dei diagrammi di persistenza con distanza di Wasserstein- p è dotato della metrica:

$$d_{Wp}(D_1, D_2) = \inf_{\{\gamma: D_1 \rightarrow D_2 \text{ biiezione parziale}\}} (\sum_{\{(b,d) \in D_1\}} \| (b,d) - \gamma(b,d) \|^{p_\infty} + \sum_{\text{punti non accoppiati diagonale}})^{1/p}$$

dove la minimizzazione è su tutte le biiezioni parziali tra i punti "off-diagonal" dei due diagrammi, con i punti non accoppiati proiettati sulla diagonale e il costo di ogni punto non accoppiato (b,d) è $\|(d-b)/2\|^p$. Per $p = \infty$ la metrica di Wasserstein- ∞ coincide con la metrica bottleneck d_B introdotta nel Capitolo 9.

La condizione di stabilità del sorites barcode (Teorema 9.2) afferma che $d_B(\text{PD}(K_1), \text{PD}(K_2)) \leq d_H(K_1, K_2)$ dove d_H è la distanza di Hausdorff tra i complessi simpliciali. Applicata alla filtrazione di Vietoris-Rips campionaria, la stabilità garantisce che $d_B(\text{PD}_n(k), \text{PD}^*(k)) \leq 2 d_H(X_n, X)$ dove $d_H(X_n, X)$ è la distanza di Hausdorff tra il campione e la distribuzione sottostante (interpretata come distanza tra la misura campionaria e la misura μ). Questa disuguaglianza, però, non è direttamente utile perché richiede di stimare $d_H(X_n, X)$, che è anch'essa una quantità stocastica.

26.2 Tassi di Convergenza della Filtrazione Campionaria

Teorema 26.1 (Convergenza del sorites barcode, Chazal et al. 2014). Sia μ una misura di probabilità su uno spazio metrico compatto (X, d) con struttura (a, b) -supporto: esiste $r_0 > 0$ tale che per ogni $x \in \text{supp}(\mu)$ e ogni $0 < r \leq r_0$ si ha $\mu(B(x, r)) \geq a \cdot r^b$. Sia $X_n = \{x_1, \dots, x_n\}$ un campione i.i.d. da μ . Allora, con probabilità almeno $1 - \delta$:

$$d_B(\text{PD}_n(k), \text{PD}^*(k)) \leq C \cdot (\log(n/\delta) / n)^{1/b}$$

dove $C > 0$ è una costante che dipende da a, b e dal diametro di X .

Questo tasso è dell'ordine $O((\log n / n)^{1/b})$, dove b è la "dimensione intrinseca" della misura μ . Per distribuzioni su spazi unidimensionali ($b = 1$), come nel caso del predicato "è alto" con $\mu =$ distribuzione delle stature nella popolazione, il tasso è $O((\log n / n))$, e la convergenza del barcode è relativamente rapida. Per distribuzioni su spazi di dimensione d maggiore (predicati multidimensionali), il tasso è più lento: $O((\log n / n)^{1/d})$. Questo rispecchia la "maledizione della dimensionalità" nella stima topologica: predicatori multidimensionali richiedono campioni molto più grandi per stimare accuratamente il loro sorites barcode.

Dimostrazione (schema). Il risultato si basa su tre passi. Passo 1: Si stima $d_H(X_n, \mu)$, la distanza di Hausdorff tra il campione X_n e il supporto di μ . Per misure con (a,b) -supporto, si dimostra che $d_H(X_n, \text{supp}(\mu)) \leq (\log(n/\delta) / (a \cdot n))^{1/b}$ con probabilità $1 - \delta$, usando una cover number argument: si copre X con ε -sfere, si stima quante celle della copertura contengono almeno un punto di X_n , e si usa la disuguaglianza di Chernoff per controllare la probabilità di celle vuote. Passo 2: Si passa dalla distanza $d_H(X_n, \text{supp}(\mu))$ alla distanza d_H tra le filtrazioni

di Vietoris-Rips $VR(X_n, r)$ e $VR(\text{supp}(\mu), r)$. Per la monotonia della filtrazione, d_H tra le filtrazioni è al più $2 \cdot d_H(X_n, \text{supp}(\mu))$. Passo 3: Si applica il teorema di stabilità del sorites barcode (Teorema 9.2) per convertire la distanza tra filtrazioni in distanza bottleneck tra diagrammi di persistenza. $\square$

Il Teorema 26.1 risponde alla domanda fondamentale con una precisione quantitativa: per avere un'approssimazione del sorites barcode con errore al più ε con probabilità almeno $1 - \delta$, è sufficiente un campione di dimensione $n \geq C' \cdot \varepsilon^{-b} \cdot \log(1/\delta)$, dove C' è una costante appropriata. Per il predicato "è alto" con $b = 1$ e $\varepsilon = 0,01$ (1% di errore nella localizzazione dei tratti del barcode) e $\delta = 0,05$ (confidenza 95%), bastano circa 500 osservazioni — un numero ragionevole per un survey sperimentale.

Proposizione 26.1 (Tasso ottimale). Il tasso $(\log n / n)^{1/b}$ del Teorema 26.1 è minimax ottimale per la stima di diagrammi di persistenza della filtrazione di Vietoris-Rips sotto la metrica bottleneck: non esiste uno stimatore che converga uniformemente più velocemente di $(1/n)^{1/b}$ (a meno di fattori logaritmici) su tutte le distribuzioni con (a,b) -supporto.

Dimostrazione. Il lower bound minimax segue dalla riduzione al problema di stima della densità in dimensione b . Si costruisce una coppia di distribuzioni μ_0, μ_1 su $X = [0,1]^b$ tali che: (i) $d_{TV}(\mu_0, \mu_1) \leq n^{-1/b}$ (distanza di variazione totale piccola, dunque i campioni di dimensione n da μ_0 e μ_1 sono statisticamente quasi indistinguibili); (ii) $d_B(\text{PD}(\mu_0), \text{PD}(\mu_1)) \geq c \cdot n^{-1/b}$ (i sorites barcode sono distinti). La costruzione è analoga a quella di Fasy et al. (2014) per la stima di diagrammi di persistenza della filtrazione di Čech: si posiziona un "buco" topologico (un ciclo in H_1) in μ_1 ma non in μ_0 , con il buco di dimensione ε , e si sceglie $\varepsilon = c \cdot n^{-1/b}$ per massimizzare la distanza dei barcode pur mantenendo la distribuzione campionaria quasi indistinguibile. $\square$

26.3 Il Bootstrap per Diagrammi di Persistenza

Il risultato del Teorema 26.1 fornisce un tasso asintotico ma non un intervallo di confidenza esplicito: la costante C non è direttamente stimabile da un campione. Il metodo bootstrap permette di costruire intervalli di confidenza nonparametrici senza conoscere C , usando la struttura del campione stesso.

Algoritmo Bootstrap per il Sorites Barcode.

Si fissi un campione X_n e si voglia costruire una banda di confidenza al livello $1 - \alpha$ per $\text{PD}_n(k)$. Il metodo è il seguente:

Per ogni $b = 1, \dots, B$ iterazioni bootstrap: si estrae un campione di risampling $X_n^{\{b\}}$ di dimensione n con rimpiazzo da X_n . Si calcola il sorites barcode $PD_n^{\{b\}}(k)$ sul campione $X_n^{\{b\}}$. Si registra la distanza bootstrap $\Delta^{\{b\}} = d_B(PD_n^{\{b\}}(k), PD_n(k))$.

Si stima il quantile bootstrap $\hat{q}_{1-\alpha}$ come il $(1-\alpha)$ -esimo quantile della distribuzione $\{\Delta^{\{b\}}\}_{b=1}^B$. La banda di confidenza bootstrap al livello $1-\alpha$ è l'intorno di $PD_n(k)$ di raggio $\hat{q}_{1-\alpha}$ nella metrica bottleneck.

Teorema 26.2 (Validità del bootstrap per PD, Chazal et al. 2014). Sotto le ipotesi del Teorema 26.1, la banda di confidenza bootstrap è asintoticamente valida:

$$\lim_{n \rightarrow \infty} P(PD^*(k) \in \{D : d_B(D, PD_n(k)) \leq \hat{q}_{1-\alpha}\}) = 1 - \alpha.$$

La validità del bootstrap si basa sul fatto che la distribuzione di $\Delta^{\{b\}} = d_B(PD_n^{\{b\}}(k), PD_n(k))$ approssima, per n grande, la distribuzione di $d_B(PD_n(k), PD^*(k))$ — l'errore reale che vogliamo quantificare. Questo segue dal principio generale di Efron che il campione bootstrap approssima la distribuzione della popolazione con errore $O(1/\sqrt{n})$ nella variazione totale, combinato con la stabilità lipschitziana del funzionale PD rispetto alla metrica di Hausdorff tra campioni.

Dimostrazione (validità del bootstrap). Per la bootstrap consistency, è sufficiente verificare che il funzionale $X_n \mapsto PD_n(k)$ soddisfi una condizione di regolarità Hadamard-differenziabilità tangenziale rispetto alla metrica di Hausdorff. Questa condizione afferma che il funzionale ha un "differenziale" D rispetto a perturbazioni della distribuzione campionaria nella direzione di processi empirici. Per i diagrammi di persistenza della filtrazione di Vietoris-Rips, la Hadamard-differenziabilità è dimostrata in Chazal, Fasy, Lecci, Rinaldo, Singh, Wasserman (2014), Theorem 2.3, usando la stabilità del diagramma di persistenza e il calcolo delle derivate di Gâteaux dei funzionali di omologia. Il teorema di validità del bootstrap di Gill e Wellner (1989) per funzionali Hadamard-differenziabili implica la validità asintotica. $\square$

Un aspetto pratico del bootstrap per il sorites barcode è la scelta del numero B di iterazioni. Per la stima del quantile al 95%, è sufficiente $B \geq 1000$ iterazioni per avere una varianza del quantile stimato inferiore a 1% del valore vero. La complessità computazionale totale dell'algoritmo bootstrap è $O(B \cdot n^2)$ per la costruzione della filtrazione (dominata dalla costruzione del grafo di tolleranza in tempo $O(n^2)$), cui si aggiunge $O(B \cdot n^3)$ per il calcolo dell'omologia persistente (dominato dalla riduzione della matrice dei bordi). Per $n \leq 500$ e $B = 1000$, il calcolo richiede tipicamente 10 – 30 minuti su hardware moderno.

26.4 Test di Ipotesi sulla Struttura di Vaghezza

La banda di confidenza bootstrap permette di formulare e testare ipotesi sulla struttura topologica del sorites barcode. Tre test statistici sono particolarmente rilevanti per le applicazioni semantiche.

Test 1: Struttura lineare vs. struttura ciclica. L'ipotesi nulla H_0 afferma che il predicato ha struttura di vaghezza lineare ($\beta_1 = 0$ nel sorites barcode, cioè nessuna barra di persistenza significativa in PD_1). L'ipotesi alternativa H_1 afferma che il predicato ha struttura ciclica (almeno una barra di lunga persistenza in PD_1). La statistica test è $T_n = \max_{(b,d) \in PD_1} (d - b)$, la persistenza massima del diagramma PD_1 calcolato sul campione. Sotto H_0 (struttura lineare), T_n è piccolo per campioni grandi (le barre di PD_1 sono tutte "rumore" con persistenza $o(n^{-1/b})$). Sotto H_1 (struttura ciclica), T_n rimane positivo al crescere di n . Il p-value del test si calcola con la bootstrap: si permuta il campione (eliminando la struttura ciclica), si ricampiona e si calcola la distribuzione di T_n^* sotto la permutazione.

Proposizione 26.2 (Consistenza del test di struttura). Il test di struttura lineare vs. ciclica con statistica T_n è consistente: sotto H_1 con $\beta_1 \geq 1$ e persistenza minima delle barre $\delta_1 > 0$, la potenza del test converge a 1 per $n \rightarrow \infty$.

Dimostrazione. Sotto H_1 , $PD^*(1)$ contiene almeno una barra (b^*, d^*) con $d^* - b^* \geq \delta_1 > 0$. Per il Teorema 26.1, $d_B(PD_n(1), PD^*(1)) \leq C(\log n / n)^{1/b} \rightarrow 0$. Dunque per n sufficientemente grande, $PD_n(1)$ contiene una barra con persistenza almeno $\delta_1 - 2C(\log n / n)^{1/b} \geq \delta_1/2$. Di conseguenza $T_n \geq \delta_1/2$ per n grande, mentre il valore critico del test (calibrato sotto H_0) converge a 0 per la consistenza del bootstrap sotto permutazione. Il test rifiuta H_0 con probabilità $\rightarrow 1$. $\square$

Test 2: Confronto di strutture di vaghezza tra predicati. Si testano due predicati P_1 e P_2 (con campioni $X_n^{\{(1)\}}$ e $X_m^{\{(2)\}}$) per verificare se hanno la stessa struttura di vaghezza. La statistica test è $T_{\{n,m\}} = d_B(PD_n^{\{(1)\}}(k), PD_m^{\{(2)\}}(k))$, la distanza bottleneck tra i barcode campionari. Il p-value si calcola con un test di permutazione: si mescolano i due campioni, si riassegnano casualmente alle due classi, e si calcola la distribuzione della statistica test sulla permutazione. Se $T_{\{n,m\}}$ è nell' $1-\alpha$ quantile della distribuzione di permutazione, i due predicati hanno strutture di vaghezza significativamente diverse al livello α .

Test 3: Variazione del sorites barcode nel tempo. Per corpora diacronici con n_t campioni per ogni anno $t = t_1 < t_2 < \dots < t_k$, si calcola la statistica $T_{\{diac\}} = \max_{s < t} d_B(PD_{\{n_s\}}(k), PD_{\{n_t\}}(k))$, la massima variazione del barcode nel tempo. Un valore alto di $T_{\{diac\}}$ indica un cambio semantico significativo nella struttura di vaghezza del predicato tra l'epoca s e l'epoca t . La distribuzione di $T_{\{diac\}}$ sotto l'ipotesi nulla di stazionarietà si stima con un bootstrap a blocchi temporali che rispetta la struttura di correlazione temporale.

26.5 Stima del Numero di Risoluzione con Intervalli di Confidenza

Il numero di risoluzione $\varepsilon^* = 1/L_{\text{mono_min}}$ è un invariante numerico del sorites barcode più semplice da stimare statisticamente rispetto all'intero diagramma di persistenza.

****Proposizione 26.3 (Stimatore di ε^*).** Sia $\varepsilon_n^* = 1/\hat{L}_n$, dove $\hat{L}_n = L_{\text{mono_min}}(G(X_n, \sim_{\{\hat{r}\}}))$ con $\hat{r}$ il sorites threshold stimato sul campione. Sotto le ipotesi del Teorema 26.1, ε_n^* è uno stimatore consistente di ε^* : $\varepsilon_n^* \rightarrow \varepsilon^*$ in probabilità per $n \rightarrow \infty$.

Dimostrazione. $L_{\text{mono_min}}$ è una funzione monotona non crescente del raggio di tolleranza r (più archi, cammini più corti). Per il Teorema 26.1, il sorites threshold stimato $\hat{r}$ soddisfa $|\hat{r} - r^*| \leq 2 \cdot d_B(\text{PD}_n(0), \text{PD}^*(0)) \rightarrow 0$ in probabilità (il sorites threshold è la persistenza del punto in PD_0 con nascita r^* che uccide l'ultima componente connessa). Per la monotonia di $L_{\text{mono_min}}$ rispetto a r , la convergenza di $\hat{r}$ a r^* implica la convergenza di $\hat{L}_n$ a $L_{\text{mono_min}}(G(X, \sim_{\{r^*\}})) = L^*$. Dunque $\varepsilon_n^* = 1/\hat{L}_n \rightarrow 1/L^* = \varepsilon^*$. $\square$

****Teorema 26.3 (Normalità asintotica di ε_n^*).** Sotto condizioni di regolarità sulla distribuzione μ (in particolare, μ ha una densità rispetto alla misura di Lebesgue e la densità è continua in un intorno del sorites threshold r^*), lo stimatore ε_n^* è asintoticamente normale:

$$\sqrt{n} \cdot (\varepsilon_n^* - \varepsilon^*) \rightarrow_d N(0, \sigma_{\varepsilon^*}^2)$$

dove $\sigma_{\varepsilon^*}^2 = (\varepsilon^*)^4 \cdot \sigma_{\{L^*\}}^2 / (L^*)^2$ e $\sigma_{\{L^*\}}^2$ è la varianza della distribuzione asintotica di $\hat{L}_n - L^*$.

La normalità asintotica permette di costruire intervalli di confidenza al livello $1-\alpha$ per ε^* come $\varepsilon_n^* \pm z_{\{\alpha/2\}} \cdot \hat{\sigma}_{\varepsilon^*} / \sqrt{n}$, dove $\hat{\sigma}_{\varepsilon^*}$ è uno stimatore bootstrap di σ_{ε^*} . Per predicatori unidimensionali con distribuzione uniforme su $[0,1]$, $\sigma_{\{L^*\}}^2 \approx (r^*)^2 / 12$, e l'intervallo di confidenza al 95% è approssimativamente $\varepsilon_n^* \pm 2 \cdot \varepsilon^* \cdot r^* / (L^* \sqrt{n})$.

Capitolo 27. Vaghezza Higher-Order: Operatori di Definiteness, Strutture Ricorsive e Geometria degli Ordini Superiori

27.1 Il Predicato Borderline e il suo Sorites Barcode

La vaghezza higher-order è il fenomeno per cui la distinzione tra casi chiari di applicazione di un predicato e casi borderline è essa stessa vaga. Non soltanto è borderline la domanda "è x alto?", ma anche la domanda "è x chiaramente alto?" può essere borderline. Questo fenomeno ha preoccupato i filosofi del linguaggio almeno da Fine (1975) e Williamson (1994), ma una trattazione matematica sistematica è rimasta elusiva. Il quadro degli spazi di tolleranza offre strumenti per formalizzare la vaghezza higher-order come una costruzione iterativa sul sorites barcode.

Definizione 27.1 (Predicato borderline). Dato uno spazio di tolleranza $(X, \sim)$ con assegnamento ε^* -ottimale $f^*: X \rightarrow [0,1]$, il predicato borderline $B(P)$ è l'insieme degli elementi $x \in X$ tali che:

(i) esiste $y \sim x$ con $f^*(y) < f^*(x)$ (x ha un vicino classificato meno chiaramente), e (ii) esiste $z \sim x$ con $f^*(z) > f^*(x)$ (x ha un vicino classificato più chiaramente).

In termini della partizione in blocchi del Teorema 3 del Capitolo 5, $B(P)$ è l'unione dei blocchi che non sono né il primo né l'ultimo nel cammino monotono minimo: $B(P) = \{x \in X : 0 < j(x) < L_mono_min\}$, dove $j(x)$ è il blocco del cammino a cui x appartiene.

La relazione di tolleranza naturale per il predicato $B(P)$ è la stessa relazione $\sim$ del predicato originale P : due elementi borderline sono indistinguibili se e solo se erano indistinguibili come istanze di P . Il sorites barcode di $B(P)$ si calcola dunque sul sottografo indotto dal grafo $G(X, \sim)$ sull'insieme $B(P)$.

Proposizione 27.1 (Sorites barcode del predicato borderline). Sia $G(B(P), \sim)$ il sottografo indotto di $G(X, \sim)$ sull'insieme borderline $B(P)$. Il numero di risoluzione di $B(P)$ soddisfa:

$$\varepsilon^*(B(P)) \leq \varepsilon^*(P) = 1/L_mono_min(X)$$

con uguaglianza se e solo se il cammino monotono minimo da $\min(B(P))$ a $\max(B(P))$ in $G(B(P), \sim)$ ha la stessa lunghezza di quello da $\min(X)$ a $\max(X)$ in $G(X, \sim)$ meno 2 passi (i passi agli estremi che escono da $B(P)$).

Dimostrazione. Il sottografo $G(B(P), \sim)$ è un sottografo di $G(X, \sim)$, dunque il cammino monotono minimo da $\min(B(P))$ a $\max(B(P))$ in $G(B(P), \sim)$ è almeno lungo quanto il corrispondente segmento del cammino minimo in $G(X, \sim)$. Il cammino minimo in $G(X, \sim)$ passa attraverso $L_mono_min + 1$ elementi, dei quali i due estremi ($\min(X)$ e $\max(X)$) non sono in $B(P)$. Il segmento interno ha $L_mono_min - 1$ archi, dunque $L_mono_min(B(P)) \geq L_mono_min(X) - 2$, e $\varepsilon^*(B(P)) = 1/L_mono_min(B(P)) \leq 1/(L_mono_min(X) - 2)$. Per L grandi, $\varepsilon^*(B(P)) \approx \varepsilon^*(P) \cdot (1 + 2/L)$. $\square$

Questa proposizione ha un'interpretazione semantica precisa: il predicato borderline è leggermente meno vago del predicato originale (ha un numero di risoluzione appena più grande), ma per predicati con lungo cammino minimo (grandi L), la differenza è trascurabile. Questo riflette l'intuizione fenomenologica che la zona borderline abbia struttura di vaghezza quasi identica al predicato originale.

27.2 L'Operatore di Definiteness e la sua Teoria Spettrale

L'operatore di definiteness graduato $\text{Def}: [0,1]^X \rightarrow [0,1]^X$ è stato introdotto nella Definizione 32 del Capitolo 17 come $\text{Def}(f)(x) = \min_{y \sim x} f(y)$. Nel presente capitolo si sviluppa la teoria spettrale di questo operatore, studiando i suoi autovettori, autovalori, e la convergenza della torre iterata $\text{Def}^k(f)$ per $k \rightarrow \infty$.

Proposizione 27.2 (Autovettori dell'operatore Def). Gli autovettori dell'operatore Def (nel senso che $\text{Def}(f) = \lambda \cdot f$ per qualche $\lambda \in \mathbb{R}$) sono esattamente le funzioni costanti su ciascuna componente connessa di $G(X, \sim)$. Per la componente connessa C_i con n_i elementi, la funzione indicatrice $1_{\{C_i\}}$ è autovettore con autovalore 1. Non esistono autovettori con autovalore diverso da 1 per l'operatore Def (poiché Def non è lineare in generale).

Per studiare la dinamica iterata, è più naturale considerare la linearizzazione dell'operatore Def intorno alla funzione ottimale f^* . Il differenziale dell'operatore Def in f^* è:

$$D(\text{Def})|_{\{f^*\}}(g)(x) = \arg \min_{y \sim x: f^*(y) = \min_{z \sim x} f^*(z)} g(y)$$

ovvero si calcola g sul vicino che realizza il minimo di f^* , e per genericità (un solo vicino realizza il minimo) questo è semplicemente la restrizione di g all'insieme dei vicini "critici" di x .

Teorema 27.1 (Convergenza dell'iterazione di Def). Per ogni assegnamento iniziale $f_0: X \rightarrow [0,1]$ ε^* -coerente, la sequenza iterata $f_k = \text{Def}^k(f_0)$ converge a una funzione limite $f_\infty: X \rightarrow [0,1]$ che è costante su ciascuna componente connessa di $G(X, \sim)$ e soddisfa $f_\infty(x) \in \{0, 1\}$ (è una classificazione bivalente). La convergenza avviene in tempo finito: esiste $K \leq |X|$ tale che $f_k = f_\infty$ per ogni $k \geq K$.

Dimostrazione. La dimostrazione procede in due passi. Passo 1 (Monotonicità): si mostra che la sequenza $f_k(x)$ è non crescente in k per ogni x . Questo segue dal fatto che $\text{Def}(f)(x) = \min_{y \sim x} f(y) \leq f(x)$ (poiché $x \sim x$ per la riflessività, dunque x è nella minimizzazione). La monotonia discendente implica la convergenza: ogni sequenza monotona discendente in $[0,1]$ converge. Passo 2 (Bivalenza del limite): si mostra che il limite f_∞ soddisfa $\text{Def}(f_\infty) = f_\infty$, cioè $f_\infty(x) = \min_{y \sim x} f_\infty(y)$ per ogni x . Questa è la condizione di "punto fisso" dell'operatore Def , e le sole funzioni che soddisfano questa condizione in ogni spazio di tolleranza connesso sono le funzioni costanti ($f_\infty \equiv c$ per qualche $c \in [0,1]$). Per componenti connesse, f_∞ è costante e il valore è determinato dal minimo degli elementi iniziali f_0 su quella componente. Poiché f_0 è ε^* -coerente con valori in $[0,1]$ e gli estremi della componente hanno valori prossimi a 0 e 1, il limite f_∞ è 0 o 1 a seconda di quale vincolo agli estremi prevale. Passo 3 (Finitezza): la convergenza è in tempo finito poiché f_k può assumere solo i valori $\{f_0(x) : x \in X\}$, un insieme finito, e la monotonicità implica che la sequenza si stabilizza dopo al più $|X|$ passi. $\square$

Il Teorema 27.1 ha un significato semantico profondo: l'applicazione iterata dell'operatore "chiaramente" su un predicato vago produce, nel limite, un predicato netto (bivalente). In altri termini, "chiaramente chiaramente ... chiaramente $P(x)$ " (con un numero sufficiente di iterazioni di "chiaramente") diventa un predicato crisp. Questo è consistente con l'intuizione di Williamson

(1994) che la vaghezza higher-order abbia profondità finita: per qualunque predicato vago finito, esiste un k^* tale che " $\text{Def}^{k^*}(P)$ " è completamente netto.

Corollario 27.1 (Profondità della vaghezza higher-order). Il minimo k^* tale che $\text{Def}^{k^*}(f_0)$ è bivalente è compreso tra 1 e $L_{\text{mono_min}} - 1$. Per la funzione ottimale $f^* = (j(x)/L)_{x \in X}$, la profondità è esattamente $k^* = L_{\text{mono_min}} - 1$: ogni applicazione di Def riduce il range effettivo di f di un passo nel cammino monotono minimo.

27.3 Teoria delle Rappresentazioni e Strutture Ricorsive

Un approccio alternativo alla vaghezza higher-order usa la teoria delle rappresentazioni di categorie. Si considera la categoria VAG i cui oggetti sono gli spazi di tolleranza $(X, \sim)$ e i cui morfismi da $(X, \sim)$ a $(Y, \approx)$ sono le funzioni $f: X \rightarrow Y$ che preservano la tolleranza ($f(x) \sim f(y)$ implica $f(x) \approx f(y)$). La vaghezza higher-order emerge come una torre di morfismi:

$$(X, \sim) \rightarrow^{\wedge \{B\}} (B(X), \sim_{\lfloor B(X) \rfloor}) \rightarrow^{\wedge \{B\}} (B^2(X), \sim_{\lfloor B^2(X) \rfloor}) \rightarrow \dots$$

dove $B^k(X) = B(B^{k-1}(X))$ è l'insieme dei casi borderline al k -esimo ordine, con la relazione di tolleranza ereditata. La torre termina al k^* -esimo livello, quando $B^{k^*}(X)$ è vuoto (tutti i casi borderline sono stati eliminati).

Definizione 27.2 (Filtrazione di borderline). La filtrazione di borderline di uno spazio di tolleranza $(X, \sim)$ è la sequenza decrescente:

$$X = B^0(X) \supseteq B^1(X) \supseteq B^2(X) \supseteq \dots \supseteq B^{k^*}(X) \supseteq B^{k^*+1}(X) = \emptyset$$

dove $B^k(X)$ è l'insieme dei casi borderline di ordine k . La profondità k^* della filtrazione è la profondità di vaghezza higher-order del predicato.

Proposizione 27.3 (Relazione tra filtrazione di borderline e cammino minimo). Per uno spazio di tolleranza $(X, \sim, \leq)$ con cammino monotono minimo $y_0 \sim y_1 \sim \dots \sim y_L$, la filtrazione di borderline soddisfa $B^k(X) \supseteq \{y_k, y_{k+1}, \dots, y_{L-k}\}$ per ogni $k \leq L/2$. In particolare, la profondità di vaghezza higher-order è $k^* = \lfloor L/2 \rfloor$.

Dimostrazione. Per $k = 0$, $B^0(X) = X$. Per $k = 1$, $B^1(X)$ contiene tutti gli elementi che non sono estremi del cammino, cioè $\{y_1, \dots, y_{L-1}\} \cup (X \setminus \{y_0, \dots, y_L\})$. Per induzione, ad ogni passo dell'iterazione il cammino si accorcia di un elemento a sinistra e uno a destra, finché non rimane vuoto dopo $\lfloor L/2 \rfloor$ iterazioni. $\square$

La Proposizione 27.3 rivela che la profondità della vaghezza higher-order di un predicato è esattamente la metà della lunghezza del suo cammino monotono minimo. Poiché $L_{\text{mono_min}} = 1/\varepsilon^*$, la profondità è $k^* \approx 1/(2\varepsilon^*)$: predicati con vaghezza piccola (ε^* vicino a 1) hanno profondità di vaghezza higher-order piccola, mentre predicati con vaghezza grande (ε^* vicino a 0) hanno profondità grande.

27.4 Vaghezza Transfinita e Punti Fissi

Per spazi di tolleranza infiniti (come il caso $X = [0,1]$ continuo o $X = \mathbb{N}$), la profondità di vaghezza higher-order può essere infinita, e la teoria della convergenza dell'iterazione di Def richiede strumenti della teoria dei punti fissi per ordini transfiniti.

Definizione 27.3 (Iterazione transfinita di Def). Per ordinali $\alpha \geq 0$, si definisce: - $f^0 = f$ (l'assegnamento iniziale) - $f^{\alpha+1} = \text{Def}(f^\alpha)$ (passo successore) - $f^\lambda = \inf_{\alpha < \lambda} f^\alpha$ (passo limite, per ordinali limite λ)

Teorema 27.2 (Convergenza transfinita). La sequenza transfinita f^α converge a un punto fisso $f^\infty = \text{Def}(f^\infty)$ per ogni assegnamento iniziale $f \in P_{\varepsilon^*}$ (l'insieme degli assegnamenti ε^* -coerenti). Il punto fisso è raggiunto all'ordinale ω (il primo ordinale infinito) per spazi di tolleranza compatti. Per spazi di tolleranza finiti, la convergenza è in tempo finito $k^* \leq |X|$ (come stabilito nel Teorema 27.1).

La dimostrazione nel caso compatto usa il Teorema di Tychonoff (il prodotto di compatti è compatto) per la convergenza di reti monotone in $[0,1]^X$, e la semicontinuità inferiore dell'operatore Def rispetto alla topologia prodotto.

Capitolo 28. Predicati Relazionali e Spazi di Tolleranza di Ordine Superiore

28.1 Vaghezza di predicati relazionali

Fino a questo punto, il saggio ha trattato predicati unari — predicati che si applicano a singoli elementi di un dominio. Tuttavia, molti predicati linguistici vaghi sono relazionali: "è simile a", "è più alto di", "somiglia a", "è vicino a". La vaghezza di questi predicati è strutturalmente diversa da quella dei predicati unari: il paradosso di Sorite per "è simile a" assume la forma "se a è simile a b e

b è simile a c , allora a è simile a c " — che è esattamente la proprietà di transitività, e la non-transitività della somiglianza è la fonte fondamentale della vaghezza semantica.

Questo capitolo sviluppa la teoria degli spazi di tolleranza di ordine superiore adatta ai predicati relazionali, definendo il sorites barcode relazionale e studiando la struttura omologica dei predicati di somiglianza.

Definizione 28.1 (Spazio di tolleranza relazionale). Dato un insieme X , un predicato relazionale binario $P: X \times X \rightarrow \{\text{Vero}, \text{Falso}\}$ (come "è simile a") è vago se la sua relazione di indistinguibilità è non transitiva. La struttura di tolleranza relazionale è una coppia $((X \times X, \sim^{\text{rel}}), \leq^{\text{rel}})$ dove:

$\sim^{\text{rel}}$ è una relazione di tolleranza su $X \times X$: $(a,b) \sim^{\text{rel}} (c,d)$ se la coppia (a,b) è "quasi indistinguibile" dalla coppia (c,d) nel senso del predicato relazionale;

$\leq^{\text{rel}}$ è un ordinamento sulle coppie: $(a,b) \leq^{\text{rel}} (c,d)$ se la coppia (a,b) è "meno simile" della coppia (c,d) nel senso del predicato.

Il grafo di tolleranza relazionale $G((X \times X), \sim^{\text{rel}})$ ha come vertici le coppie ordinate $(a,b) \in X \times X$ e archi tra coppie $\sim^{\text{rel}}$ -tolleranti. Il complesso di Sorite relazionale $S((X \times X), \sim^{\text{rel}})$ è il complesso delle clique di questo grafo.

Esempio 28.1 (Predicato "è simile a" per colori). Si consideri $X = \{\text{rosso}, \text{arancio}, \text{giallo}, \text{verde}, \text{blu}, \text{viola}\}$ (6 colori fondamentali) con il predicato "è simile a" inteso come adiacenza sullo spettro. Due colori sono indistinguibili se sono adiacenti sullo spettro. Le coppie (rosso, arancio) e (arancio, giallo) sono $\sim^{\text{rel}}$ -tolleranti perché la differenza di un "passo" nella catena di colori è la stessa. Le coppie (rosso, arancio) e (giallo, verde) non sono $\sim^{\text{rel}}$ -tolleranti perché si trovano in zone diverse dello spettro. Il grafo $G((X \times X), \sim^{\text{rel}})$ ha 36 vertici (le 6^2 coppie) e archi che riflettono la struttura dell'ordinamento spettrale.

28.2 Il Sorites Relazionale e la Non-Transitività della Somiglianza

Il paradosso di Sorite per i predicati relazionali ha la struttura seguente:

P1: "L'arancio è simile al rosso" (caso paradigmatico di somiglianza). P2: Se il colore c_1 è simile al colore c_2 e il colore c_2 è simile al colore c_3 , allora c_1 è simile a c_3 (principio di tolleranza per la somiglianza, equivalente alla transitività). C: "Il rosso è simile al verde" (per una catena di somiglianze: rosso $\sim$ arancio $\sim$ giallo $\sim$ verde).

La conclusione è ovviamente falsa: rosso e verde non sono simili. La struttura logica è identica al sorite classico, con P2 che richiede la transitività della somiglianza. La risoluzione formale è analoga: P2 deve essere indebolito, e il "costo" di questo indebolimento è misurato dal numero di risoluzione del predicato relazionale.

Definizione 28.2 (Numero di risoluzione relazionale). Il numero di risoluzione del predicato relazionale "è simile a" su X è:

$\varepsilon_{\sim}^{\text{rel}}(\sim) = \min\{\varepsilon \geq 0 : \text{esiste una funzione di membership relazionale } f: X \times X \rightarrow [0,1] \text{ che è } \varepsilon\text{-coerente, monotona rispetto a } \leq^{\text{rel}}, \text{ e soddisfa } f(a,a) = 1 \text{ (riflessività) e } f(a,b) = f(b,a) \text{ (simmetria) per ogni } a,b \in X\}$

Un assegnamento f è ε -coerente nel senso relazionale se $|f(a,b) - f(c,d)| \leq \varepsilon$ ogni volta che $(a,b) \sim^{\text{rel}} (c,d)$.

Proposizione 28.1 (Connessione con il numero di risoluzione unario). Sia P un predicato unario vago su $(X, \sim, \leq)$. Il predicato relazionale "x è meno P di y" (o equivalentemente "y è più P di x") ha struttura di tolleranza $\sim^{\text{rel}}$ su $X \times X$ definita da $(x,y) \sim^{\text{rel}} (x',y')$ se $x \sim x'$ e $y \sim y'$, con ordinamento $(x,y) \leq^{\text{rel}} (x',y')$ se $x \leq x'$ e $y \geq y'$ (l'"ascesa" nella scala del predicato). Il numero di risoluzione relazionale soddisfa:

$$\varepsilon_{\sim}^{\text{rel}}(\sim) \leq \varepsilon^*(\sim) + \varepsilon^*(\sim) = 2\varepsilon^*(\sim)$$

con uguaglianza per predicati con struttura di catena lineare.

Dimostrazione. La funzione $f(x,y) = \max(0, f^*(y) - f^*(x))$ (dove $f^*: X \rightarrow [0,1]$ è la funzione ottimale del predicato unario) è una funzione di membership relazionale. Per $x \sim x'$ e $y \sim y'$, si ha:

$$|f(x,y) - f(x',y')| = |\max(0, f^*(y) - f^*(x)) - \max(0, f^*(y') - f^*(x'))|$$

$$\leq |f^*(y) - f^*(y')| + |f^*(x) - f^*(x')| \leq \varepsilon^*(\sim) + \varepsilon^*(\sim) = 2\varepsilon^*(\sim)$$

poiché $\max(0, \cdot)$ è 1-Lipschitz e f^* è ε^* -coerente. $\square$

28.3 Omologia del Grafo di Tolleranza Relazionale

Il grafo di tolleranza relazionale $G((X \times X), \sim^{\text{rel}})$ ha una struttura molto più ricca di quella del grafo unario $G(X, \sim)$, e la sua omologia può avere componenti in gradi superiori anche per predicati linguistici naturali.

Teorema 28.1 (Omologia del grafo di tolleranza per la somiglianza circolare). Per il predicato "è simile a" su $X = \{1, 2, \dots, n\}$ con somiglianza circolare: $i \sim^{\text{rel}} j$ se $|i - j| \leq 1 \pmod{n}$, e $(i, j) \sim^{\text{rel}} (k, l)$ se $|i - k| \leq 1$ e $|j - l| \leq 1 \pmod{n}$, il grafo di tolleranza relazionale $G((X \times X), \sim^{\text{rel}})$ ha:

$\beta_0(G) = 1$ (il grafo è connesso per $n \geq 3$) $\beta_1(G) = 2$ per $n \geq 4$ (due cicli indipendenti)

I due cicli indipendenti corrispondono ai due loop fondamentali della struttura toroidale $X \times X$ (cicli in ciascuna delle due coordinate).

Dimostrazione. Il grafo $G((X \times X), \sim^{\text{rel}})$ è isomorfo al prodotto cartesiano dei grafi $C_n \times C_n$, dove C_n è il ciclo di lunghezza n . Per la formula di Künneth per l'omologia del prodotto:

$$H_k(C_n \times C_n; \mathbb{Z}) \cong \bigoplus_{i+j=k} (H_i(C_n; \mathbb{Z}) \otimes H_j(C_n; \mathbb{Z})) \oplus \bigoplus_{i+j=k-1} \text{Tor}(H_i(C_n; \mathbb{Z}), H_j(C_n; \mathbb{Z}))$$

Per C_n : $H_0(C_n) = \mathbb{Z}$, $H_1(C_n) = \mathbb{Z}$, $H_k(C_n) = 0$ per $k \geq 2$. Dunque:

$$H_0(C_n \times C_n) = \mathbb{Z} \otimes \mathbb{Z} = \mathbb{Z} \rightarrow \beta_0 = 1 \quad H_1(C_n \times C_n) = (\mathbb{Z} \otimes \mathbb{Z}) \oplus (\mathbb{Z} \otimes \mathbb{Z}) = \mathbb{Z}^2 \rightarrow \beta_1 = 2$$

I due generatori di H_1 sono il ciclo in "direzione x" (tenendo y fissa) e il ciclo in "direzione y" (tenendo x fissa). $\square$

Questo teorema mostra che i predicati di somiglianza circolare — come la somiglianza cromatica su uno spazio di colori circolare — hanno genericamente $\beta_1 = 2$, in contrasto con i predicati unari lineari ($\beta_1 = 0$) o circolari ($\beta_1 = 1$). La struttura toroidale della somiglianza produce ostruzioni omologiche di ordine superiore: esistono due cicli indipendenti nel grafo di tolleranza relazionale, corrispondenti a due paradossi di Sorite indipendenti e non riducibili l'uno all'altro.

28.4 Predicati Comparativi e la Struttura degli Semiordini

Molti predicati relazionali vaghi sono comparativi: "è più alto di", "è più veloce di", "è più calvo di". La struttura formale di questi predicati è quella dei semiordini (semiorders), introdotti da Luce (1956) per modellare le preferenze con soglie percettive.

Definizione 28.3 (Semiordine). Un semiordine su X è una relazione binaria $>$ che soddisfa: (S1) irreflessive: $\neg(x > x)$; (S2) transitività debole: $x > y$ e $z > w$ implica $x > w$ o $z > y$; (S3) no due catene in parallelo: $\neg(x > y > z)$ implica $\neg(x > w)$ o $\neg(w > z)$ per ogni w .

Teorema 28.2 (Scott-Suppes, 1958). Una relazione $>$ su un insieme numerabile X è un semiordine se e solo se esiste una funzione $u: X \rightarrow \mathbb{R}$ e un threshold $\delta > 0$ tali che $x > y \leftrightarrow u(x) > u(y) + \delta$.

La rappresentazione di Scott-Suppes collega i semiordini alla teoria degli spazi di tolleranza: la relazione di indifferenza (non preferenza in nessun senso) $x \sim y \leftrightarrow \neg(x > y) \text{ e } \neg(y > x)$ è esattamente la relazione di tolleranza con soglia δ . La funzione u è la funzione di membership del predicato comparativo.

Proposizione 28.2 (Numero di risoluzione di un semiordine). Per un semiordine $(X, >)$ con rappresentazione Scott-Suppes (u, δ) e dominio $X = \{x_1, \dots, x_n\}$ con $u(x_1) < u(x_2) < \dots < u(x_n)$, il numero di risoluzione del predicato comparativo " $x > y$ " è:

$$\varepsilon^*_{\text{comp}} = \delta / (u(x_n) - u(x_1))$$

che è il rapporto tra la soglia percettiva δ e l'ampiezza del range di utilità.

Capitolo 29. Variazione Diacronica della Vaghezza: Persistenza Zigzag e Topologia del Cambiamento Semantico

29.1 Filtrazioni Variabili nel Tempo e Persistenza Zigzag

La struttura di vaghezza di un predicato non è immutabile nel tempo: le norme linguistiche evolvono, le soglie di indistinguibilità si spostano con i mutamenti culturali, tecnologici e demografici. Un esempio emblematico è il predicato "è giovane": nel XIX secolo, "giovane" si applicava tipicamente a individui sotto i 30 anni; oggi, la stessa parola viene comunemente usata per individui fino ai 40-45 anni, riflettendo il mutamento delle aspettative di vita e delle strutture sociali. Il sorites threshold r^* del predicato "è giovane" si è quindi spostato verso destra nel corso del tempo, e il corrispondente numero di risoluzione ε^* è cambiato.

Per trattare matematicamente la variazione diacronica del sorites barcode, si usa la teoria della persistenza zigzag, sviluppata da Carlsson e de Silva (2010) per studiare sequenze di spazi topologici collegati da mappe in entrambe le direzioni.

Definizione 29.1 (Filtrazione zigzag del predicato diacronico). Sia $\{(X, \sim_{\{t\}})\}_{t \in T}$ una famiglia di spazi di tolleranza parametrizzata dal tempo $t \in T = \{t_1 < t_2 < \dots < t_k\}$, dove $\sim_{\{t_i\}}$ è la relazione di tolleranza del predicato al tempo t_i (estratta da corpora di testi dell'epoca t_i). La filtrazione zigzag associata è la sequenza di complessi e mappe:

$$S(X, \sim_{\{t_1\}}) \leftarrow^{\{f_1\}} S(X, \sim_{\{t_1\}} \cup \sim_{\{t_2\}}) \rightarrow^{\{g_1\}} S(X, \sim_{\{t_2\}}) \leftarrow^{\{f_2\}} S(X, \sim_{\{t_2\}} \cup \sim_{\{t_3\}}) \rightarrow^{\{g_2\}} S(X, \sim_{\{t_3\}}) \leftarrow \dots \rightarrow S(X, \sim_{\{t_k\}})$$

dove $\sim_{\{t_i\}} \cup \sim_{\{t_{i+1}\}}$ è la relazione di tolleranza "amalgamata" che combina quelle dei due periodi consecutivi, e f_i, g_i sono le inclusioni (ogni complesso nel periodo è un sottocomplesso dell'amalgama).

Teorema 29.1 (Teorema di struttura per la persistenza zigzag, Carlsson-de Silva 2010). Ogni modulo zigzag:

$$M_1 \leftarrow^{\{f_1\}} M_2 \rightarrow^{\{g_1\}} M_3 \leftarrow M_4 \rightarrow \dots \leftarrow M_{2k-1} \rightarrow M_{2k}$$

(dove M_i sono spazi vettoriali su un campo k e f_i, g_i sono applicazioni lineari) si decompone come somma diretta di moduli intervallo zigzag:

$$M \cong \bigoplus_j I[a_j, b_j]$$

dove $I[a,b]$ è il modulo zigzag che vale k sulle posizioni da a a b e 0 altrove (con mappe identità nelle posizioni interne). Il multiinsieme $\{[a_j, b_j]\}$ è l'invariante completo del modulo e generalizza il diagramma di persistenza al caso zigzag.

Il diagramma di persistenza zigzag del sorites barcode diacronico è dunque ben definito e classifica completamente la variazione temporale della struttura di vaghezza del predicato. Ogni barra $[a,b]$ nel diagramma indica un "tratto topologico" (una componente connessa, un ciclo omologico) che è presente durante il periodo da t_a a t_b e assente fuori da questo intervallo.

29.2 Cambiamento Semantico come Deformazione Topologica

Il quadro della persistenza zigzag permette di distinguere tre tipi qualitativamente diversi di cambiamento semantico diacronico, in termini della struttura dei diagrammi zigzag:

Tipo I — Deriva graduale: il sorites threshold $r^*(t)$ varia in modo continuo nel tempo, e il diagramma di persistenza zigzag ha barre di lunga durata che corrispondono a tratti strutturali persistenti (stessa struttura topologica, valori quantitativi spostati). Questo è il caso del predicato "è giovane" menzionato sopra.

Tipo II — Biforcazione semantica: un predicato si "sdoppia" in due precisazioni con strutture di vaghezza diverse. Nel diagramma zigzag, una barra in PD_0 scompare e due nuove barre compaiono. Un esempio storico è il predicato "è dotto" in italiano, che fino al XVIII secolo aveva un'unica zona di applicazione (persone con conoscenza accademica), e nel XIX si è biforcato in "è erudito"

(conoscenza storica e letteraria) e "è scientifico" (conoscenza sperimentale), con strutture di vaghezza distinte.

Tipo III — Rottura semantica: il sorites barcode cambia discontinuamente, segnalando una riorganizzazione radicale della struttura di vaghezza. Nel diagramma zigzag, tutte le barre vecchie si chiudono in un breve intervallo e nuove barre emergono. Questo corrisponde a quello che i linguisti chiamano "semantic bleaching" o "narrowing" — fenomeni come il passaggio dell'italiano "comodo" da "accessibile" (di percorso) a "confortevole" (di oggetto) avvenuto tra il XVI e il XVIII secolo.

Teorema 29.2 (Invarianza del numero di barre per derive graduate). Se la variazione diacronica è di Tipo I — la distanza bottleneck $d_B(\text{PD}_n(X, \sim_{\{t_i\}}), \text{PD}_n(X, \sim_{\{t_{i+1}\}})) \leq \varepsilon$ per ogni passo temporale i — allora il numero di barre di persistenza nel diagramma zigzag al livello di persistenza $> 2\varepsilon$ è costante nel tempo e uguale al numero di barre del barcode iniziale $\text{PD}(X, \sim_{\{t_1\}})$ al livello > 0 .

Dimostrazione. Questo segue dalla stabilità della persistenza zigzag: per il teorema di stabilità dei diagrammi zigzag (Botnan-Lesnick 2018), $d_B(\text{PD}^{\{\text{zig}\}}(M), \text{PD}^{\{\text{zig}\}}(N)) \leq \|M - N\|_{\{\text{op}\}}$, dove $\|M - N\|_{\{\text{op}\}}$ è la norma operatoriale della differenza dei moduli zigzag. Per una sequenza con variazioni gradualmente $\leq \varepsilon$ ad ogni passo, la norma operatoriale è al più ε per costruzione. Il numero di barre con persistenza $> 2\varepsilon$ è dunque invariante per perturbazioni di norma $\leq \varepsilon$. $\square$

29.3 La Distanza Diacronica e il Cambiamento Semantico Misurato

Per quantificare il cambiamento semantico complessivo tra due epoche t_1 e t_2 , si introduce la distanza diacronica del sorites barcode.

Definizione 29.2 (Distanza diacronica). Per due epoche $t_1 < t_2$ e per il k -esimo grado omologico, la distanza diacronica è:

$$d_{\{\text{diac}\}}(t_1, t_2; k) = d_B(\text{PD}_k(X, \sim_{\{t_1\}}), \text{PD}_k(X, \sim_{\{t_2\}}))$$

ovvero la distanza bottleneck tra i sorites barcode delle due epoche. Una distanza diacronica piccola indica che la struttura di vaghezza è rimasta sostanzialmente invariata; una distanza grande indica un cambiamento semantico significativo.

Proposizione 29.1 (Subadditività della distanza diacronica). Per tre epoche $t_1 < t_2 < t_3$:

$$d_{\{\text{diac}\}}(t_1, t_3; k) \leq d_{\{\text{diac}\}}(t_1, t_2; k) + d_{\{\text{diac}\}}(t_2, t_3; k)$$

La disuguaglianza triangolare è la proprietà naturale di una "distanza semantica" nel tempo: il cambiamento complessivo da t_1 a t_3 è al più la somma dei cambiamenti intermedi.

Dimostrazione. Per le proprietà metriche della distanza bottleneck: $d_B(PD_1, PD_3) \leq d_B(PD_1, PD_2) + d_B(PD_2, PD_3)$ per la disuguaglianza triangolare di d_B . $\square$

Questa struttura metrica sullo spazio dei predicati $\times$ epoche permette di costruire una "mappa semantica diacronica": ciascun predicato in ciascuna epoca è un punto in uno spazio metrico, e le distanze tra punti misurano il grado di cambiamento semantico. La tecnica del multidimensional scaling applicata a questa matrice di distanze produce una visualizzazione a bassa dimensione del "paesaggio diacronico" della vaghezza semantica.

29.4 Analisi di Corpora Storici: Metodologia e Risultati Attesi

La teoria sviluppata nelle sezioni precedenti suggerisce una metodologia concreta per l'analisi diacronica della vaghezza usando corpora di testi storici come Google Books Ngrams, il Corpus dell'Italiano scritto (CIS), o il corpus CORIS per l'italiano contemporaneo.

La metodologia si articola come segue. Per ciascuna epoca t_i e ciascun predicato P in esame, si estraggono dal corpus le istanze con qualificatori di vaghezza: "chiaramente P ", "abbastanza P ", "quasi P ", "non esattamente P ", "borderline P ". Si calcola la frequenza di co-occorrenza di ciascuna istanza con ciascun qualificatore, normalizzata per la frequenza totale dell'istanza. Si definisce una distanza tra istanze basata sulla divergenza di Kullback-Leibler delle distribuzioni di qualificatori: $d_{\{KL\}}(x, y) = \sum_q P_x(q) \cdot \log(P_x(q)/P_y(q))$, dove $P_x(q)$ è la probabilità che il qualificatore q co-occorra con x nel corpus. Si costruisce la filtrazione di Vietoris-Rips con la metrica $d_{\{KL\}}$ e si calcola il sorites barcode per ogni epoca.

Per il predicato "è ricco" in italiano tra il 1850 e il 2020 (usando Google Ngrams), la previsione della teoria è che il sorites threshold $r^*(t)$ sia aumentato nel tempo (le soglie percettive di "ricchezza" si sono spostate verso l'alto in termini di potere d'acquisto), ma che il numero di risoluzione $\epsilon^*(t) = 1/L_{\text{mono_min}}(t)$ sia rimasto sostanzialmente costante (la struttura relativa della vaghezza è invariante anche se i valori assoluti cambiano). Questa previsione è empiricamente verificabile con i dati di corpus disponibili, e costituisce un test non banale della teoria degli spazi di tolleranza applicata alla linguistica storica.

Capitolo 30. Apprendimento Automatico della Struttura di Tolleranza: Dal Problema Inverso all'Inferenza Topologica

30.1 Il Problema Inverso

Il problema inverso della teoria degli spazi di tolleranza è il seguente: dati N osservazioni di classificazione $D = \{(x_i, y_i)\}_{i=1}^N$, dove $x_i \in X$ è un'istanza del predicato e $y_i \in \{0, 1\}$ è il giudizio di classificazione emesso da un parlante, determinare la struttura di tolleranza $\sim$ su X che meglio "spiega" i dati osservati.

La difficoltà del problema è che la struttura di tolleranza non è direttamente osservabile: si osservano classificazioni bivalenti (0 o 1), non relazioni di indistinguibilità. La struttura di tolleranza deve essere inferita indirettamente, osservando quali coppie (x_i, x_j) ricevono classificazioni discordanti tra parlanti diversi o in contesti diversi.

Definizione 30.1 (Funzione di perdita per l'apprendimento della tolleranza). Data una struttura di tolleranza $\sim$ su X e un campione di classificazioni D , la funzione di perdita ε -inconsistente è:

$$L_{\varepsilon}(\sim, D) = (1/N) \sum_{i=1}^N \max(0, \min_{y \sim x_i} y_i + \varepsilon - \max_{y \sim x_i} y_i) + \lambda \cdot C(\sim)$$

dove $C(\sim)$ è un termine di regolarizzazione che penalizza la complessità della struttura di tolleranza (ad esempio, il numero di archi del grafo $G(X, \sim)$), e $\lambda \geq 0$ è un parametro di regolarizzazione. Il termine di perdita $\max(0, \dots)$ è positivo quando la variazione delle classificazioni nella stella $St(x_i)$ supera ε : in tal caso, non esiste un assegnamento ε -coerente che classifichi correttamente x_i .

Il problema di apprendimento è dunque: minimizzare $L_{\varepsilon}(\sim, D)$ su tutte le strutture di tolleranza ammissibili (relazioni simmetriche e riflessive su X), trovando la struttura $\sim$ che minimizza l'inconsistenza ε con i dati osservati, penalizzata per la complessità.

30.2 Convessità e Rilassamenti

Il problema di ottimizzazione su strutture di tolleranza è combinatoriale: ci sono $2^{\binom{n-1}{2}}$ possibili strutture di tolleranza su un insieme di n elementi. Per $n = 100$, questo è un numero astronomico che rende l'ottimizzazione esatta computazionalmente impossibile. Si considerano quindi rilassamenti convessi.

Rilassamento 1: Variabili continue di adiacenza. Si sostituisce la variabile binaria $1_{\{x \sim y\}} \in \{0,1\}$ con una variabile continua $w_{\{xy\}} \in [0,1]$, che rappresenta il "grado" di adiacenza tra x e y . La funzione di perdita rilassata è:

$$\tilde{L}_\varepsilon(W, D) = (1/N) \sum_{i=1}^N \max(0, \text{var}(y_i | \text{St}(x_i)^W) - \varepsilon^2) + \lambda \cdot \|W\|_1$$

dove $\text{St}(x_i)^W$ è la stella "soft" pesata da W e $\|W\|_1 = \sum_{x,y} w_{xy}$ è la regolarizzazione L^1 che promuove la sparsità (pochi archi).

Questo rilassamento è convesso nel parametro W (il max di funzioni lineari è convesso, e la norma L^1 è convessa), dunque si può ottimizzare con algoritmi di programmazione convessa in tempo polinomiale.

Proposizione 30.1 (Convessità del rilassamento). La funzione $\tilde{L}_\varepsilon(W, D)$ è convessa in W per ogni fissato $\varepsilon \geq 0$. L'ottimizzatore W^* esiste ed è unico se $\lambda > 0$.

Dimostrazione. La convessità segue dall'osservazione che $\max(0, f(W))$ è convessa se f è convessa o affine. $\text{var}(y_i | \text{St}(x_i)^W)$ è un'espressione quadratica nelle y_i pesata da w_{xy} ; la composizione con $\max(0, \cdot)$ preserva la convessità. La norma L^1 è convessa. La somma di funzioni convesse è convessa. L'unicità per $\lambda > 0$ segue dalla stretta convessità della regolarizzazione L^1 rinforzata. $\square$

Rilassamento 2: Strutture di tolleranza spettrali. Un'altra strada di rilassamento usa la rappresentazione spettrale del grafo di tolleranza. Invece di ottimizzare direttamente sugli archi, si ottimizzano i $k \leq n$ autovettori del laplaciano normalizzato $L(\sim)$, cercando la struttura di autovettori che minimizza l'inconsistenza con i dati. Questo approccio è collegato alla "spectral embedding" usata nel clustering spettrale (Ng, Jordan, Weiss 2002).

30.3 Bound di Generalizzazione Topologici

Una domanda fondamentale nell'apprendimento automatico è la capacità di generalizzazione: quanto bene la struttura di tolleranza appresa su un campione training prevede le classificazioni su nuovi dati non visti? I classici risultati di VC theory (Vapnik-Chervonenkis 1974) forniscono bound sul rischio atteso in termini della dimensione VC della classe di ipotesi. Per la classe delle strutture di tolleranza, la dimensione VC ha una caratterizzazione topologica precisa.

Definizione 30.2 (Dimensione VC della classe di strutture di tolleranza). La dimensione VC della classe $T_n = \{\sim: \text{struttura di tolleranza su } X \text{ con } |X| = n\}$ rispetto alla classe di classificatori indotti $\{f_\sim: X \rightarrow \{0,1\} \text{ coerente con } \sim\}$ è:

$\text{VC-dim}(T_n) = \max\{|S| : S \subseteq X, \text{ogni subset di } S \text{ è classificabile da qualche } f_\sim \text{ coerente con qualche } \sim \in T_n\}$

Teorema 30.1 (Bound sulla dimensione VC). Per strutture di tolleranza su X con $|X| = n$:

$$\lfloor \log_2 n \rfloor \leq \text{VC-dim}(T_n) \leq n - 1$$

Il lower bound $\lfloor \log_2 n \rfloor$ è raggiunto da strutture di tolleranza con grafo completo bipartito (ogni subset di dimensione $\leq \log_2 n$ è separabile). Il upper bound $n-1$ è raggiunto da strutture di tolleranza con grafo catena (ogni subset è separabile purché abbia la struttura ordinale corretta).

Dimostrazione (lower bound). Si sceglie $S \subseteq X$ con $|S| = \lfloor \log_2 n \rfloor$. Per ogni sottoinsieme $A \subseteq S$, si costruisce una struttura di tolleranza $\sim_A$ con $k = \lfloor n/|S| \rfloor$ componenti connesse, dove ciascun elemento di A è in una componente "vero" e ciascun elemento di $S \setminus A$ è in una componente "falso". Il classificatore bivalente che assegna "vero" alle componenti di A e "falso" alle componenti di $S \setminus A$ classifica correttamente tutti gli elementi di S . Poiché ci sono $2^{|S|} = 2^{\lfloor \log_2 n \rfloor} \leq n$ sottoinsiemi di S , esiste una struttura di tolleranza per ciascuno di essi. Dunque S è frantumatamente da T_n , e $\text{VC-dim}(T_n) \geq \lfloor \log_2 n \rfloor$. $\square$

Corollario 30.1 (Bound di generalizzazione per strutture di tolleranza). Con N campioni di addestramento e probabilità almeno $1 - \delta$, il rischio empirico-to-vero dell'apprendimento di strutture di tolleranza soddisfa:

$$R(\sim_N) \leq \hat{R}_N(\sim_N) + O(\sqrt{(\text{VC-dim}(T_n) \cdot \log(N/\delta) / N)}) = \hat{R}_N(\sim_N) + O(\sqrt{(n \cdot \log(N/\delta) / N)})$$

Per $n = 100$ istanze del predicato e $N = 10000$ esempi di addestramento con $\delta = 0,05$, il bound è $R \leq \hat{R} + O(0,05)$, che è accettabile per applicazioni pratiche.

30.4 Apprendimento Online e Aggiornamento del Sorites Barcode

In molte applicazioni pratiche, i dati di classificazione arrivano in modo sequenziale nel tempo: ogni nuovo giudizio di un parlante aggiorna le stime della struttura di tolleranza. Il framework online dell'apprendimento richiede algoritmi che aggiornino la struttura di tolleranza in modo efficiente ad ogni nuovo campione, senza dover ricalcolare l'intera struttura da zero.

Algoritmo 30.1 (Aggiornamento online della struttura di tolleranza). Si parte da una struttura di tolleranza iniziale $G_0 = G(X, \sim_0)$ e si ricevono sequenzialmente classificazioni (x_t, y_t) per $t = 1, 2, \dots$. Per ogni nuovo campione (x_t, y_t) , si aggiorna la struttura come segue:

1. Si calcola il voto di consenso $\hat{C}(x_t) = \text{round}((1/|St(x_t)|) \sum \{y \sim x_t\} y_t)$, dove la media è sui vicini già classificati.
2. Se $\hat{C}(x_t) = y_t$ (consensus), non si modifica la struttura.
3. Se $\hat{C}(x_t) \neq y_t$ (dissensus), si cerca l'arco $\{x_t, z_t\}$ con z_t classificato diversamente da x_t che crea maggiore inconsistenza, e si rimuove quell'arco dalla struttura (si indebolisce la relazione di tolleranza).

Proposizione 30.2 (Bound sul numero di errori online). Il numero di errori dell'algoritmo di aggiornamento online (passi in cui il consensus è sbagliato) è al più $O(\varepsilon^* \cdot N)$, dove ε^* è il numero di risoluzione della struttura di tolleranza vera e N è il numero totale di campioni.

Dimostrazione (schema). La prova usa un argomento di "margin": la struttura di tolleranza "vera" $\sim^*$ ha ε^* come numero di risoluzione, e ogni classificazione inconsistente riduce il margine di ε^* di almeno $1/N$. Poiché il margine totale disponibile è ε^* (dalla struttura ottimale), il numero totale di riduzioni è al più $\varepsilon^* \cdot N$. $\square$

Questo bound mostra che l'apprendimento online della struttura di tolleranza è efficiente: il numero di errori nel tempo è lineare in $\varepsilon^* \cdot N$, e per predicati con vaghezza piccola (ε^* piccolo), il numero di errori è anch'esso piccolo. In modo controintuitivo, i predicati meno vaghi (ε^* vicino a 1) sono più difficili da apprendere online: la struttura di tolleranza quasi crisp richiede molti esempi discordanti per individuare i confini.

Capitolo 31. Geometria Discreta della Vaghezza: Tra Combinatoria e Analisi

31.1 Il Laplaciano di Hodge per Complessi di Sorite

La teoria spettrale del Capitolo 8 ha sviluppato il laplaciano del grafo come operatore sul dominio 0-dimensionale (funzioni sui vertici). Il laplaciano di Hodge generalizza questo operatore a tutte le dimensioni, fornendo un'analisi spettrale completa del complesso di Sorite.

Definizione 31.1 (Laplaciano di Hodge). Per un complesso simpliciale K con complessi di catene $C_k(K; \mathbb{R})$ e operatori di bordo $\partial_k: C_k \rightarrow C_{k-1}$, si definisce l'operatore aggiunto $\partial_k^*: C_{k-1} \rightarrow C_k$ come il trasposto di ∂_k rispetto al prodotto scalare standard. Il k -esimo laplaciano di Hodge è:

$$\Delta_k = \partial_{k+1} \partial_{k+1}^* + \partial_k^* \partial_k: C_k(K; \mathbb{R}) \rightarrow C_k(K; \mathbb{R})$$

Per $k = 0$, si ottiene $\Delta_0 = \partial_1 \partial_1^*$ (il laplaciano del grafo, poiché $\partial_0 = 0$). Per $k = 1$, $\Delta_1 = \partial_2 \partial_2^* + \partial_1^* \partial_1$ è il laplaciano sugli archi.

Teorema di Hodge discreto. Il nucleo di Δ_k è isomorfo al k -esimo gruppo di omologia reale: $\ker(\Delta_k) \cong H_k(K; \mathbb{R})$. In particolare, la dimensione del nucleo di Δ_k è il k -esimo numero di Betti β_k .

Dimostrazione. Per la decomposizione di Hodge, $C_k(K; \mathbb{R}) = \ker(\Delta_k) \oplus \text{im}(\partial_{k+1}) \oplus \text{im}(\partial_k^*)$. Questa decomposizione è ortogonale rispetto al prodotto scalare standard. Il nucleo di Δ_k è l'insieme delle k -catene che sono simultaneamente cicli ($\ker \partial_k$) e cocicli ($\ker \partial_k^*$), cioè le k -forme armoniche. Il teorema dei coefficienti universali implica $\ker(\Delta_k) \cong H_k(K; \mathbb{R}) = H_k(K; \mathbb{Z}) \otimes \mathbb{R}$ per complessi con omologia libera. $\square$

Per il complesso di Sorite $S(X, \sim)$, il laplaciano di Hodge Δ_1 agisce sugli archi (1-simplessi) del grafo di tolleranza. Gli autovettori di Δ_1 con autovalore 0 sono i cicli armonici — le combinazioni lineari di archi che sono sia cicli (bordo zero) che cocicli (coogradiente zero). Il numero di tali cicli armonici è $\beta_1 = \text{rk}(H_1(K; \mathbb{Z}))$, il numero di ostruzioni omologiche del predicato.

Proposizione 31.1 (Spectrum di Δ_1 per il ciclo C_n). Per il complesso di Sorite del ciclo C_n (il grafo di Sorite ciclico), gli autovalori di Δ_1 sono:

$$\lambda_k = 2 - 2\cos(2\pi k/n) \text{ per } k = 0, 1, \dots, n-1$$

con autovettori $\psi_k(j) = e^{2\pi i j k/n}$ (le onde di Fourier discrete). L'autovalore $\lambda_0 = 0$ ha molteplicità 1 e corrisponde all'unico ciclo armonico $\beta_1 = 1$ del grafo C_n .

Dimostrazione. Δ_1 per il ciclo C_n è la matrice di convoluzione circolare $\tilde{L} = L \otimes I_n - A_1$ (dove L è il laplaciano del grafo e A_1 è la matrice di adiacenza agisce sugli archi). Per simmetria circolare, gli autovettori sono le onde di Fourier $e^{2\pi i j k/n}$, e gli autovalori si calcolano come la trasformata di Fourier discreta della sequenza di convoluzione $(2, -1, 0, \dots, 0, -1)$: $\lambda_k = 2 - 2\cos(2\pi k/n)$. $\square$

31.2 La Teoria di Morse Discreta per i Complessi di Sorite

La teoria di Morse discreta, sviluppata da Forman (1998), fornisce un'altra prospettiva sulla topologia del complesso di Sorite. Una funzione di Morse discreta su un complesso K è un'assegnazione $f: K \rightarrow \mathbb{R}$ (dove K è l'insieme di tutti i simplessi di K) tale che per ogni k -simpleso σ :

- (i) al più un $(k+1)$ -simpleso $\tau \supset \sigma$ soddisfa $f(\tau) \leq f(\sigma)$; (ii) al più un $(k-1)$ -simpleso $\tau \subset \sigma$ soddisfa $f(\tau) \geq f(\sigma)$.

Un semplice σ è critico se non soddisfa nessuna delle relazioni di accoppiamento descritte in (i) e (ii). Il Teorema di Morse di Forman afferma che il complesso K ha lo stesso tipo di omotopia di un CW-complesso con una cella di dimensione k per ogni semplice critico di dimensione k .

Proposizione 31.2 (Funzione di Morse ottimale per il complesso di Sorite). Per il complesso di Sorite $S(X, \sim)$ di un predicato lineare (grafo di tolleranza che è un albero o un grafo aciclico), esiste una funzione di Morse discreta con esattamente β_0 semplici critici di dimensione 0, β_1 semplici critici di dimensione 1, e nessun semplice critico di dimensione superiore.

Dimostrazione. Si costruisce la funzione di Morse usando la funzione ottimale $f^*: X \rightarrow [0,1]$ del Teorema 3. Si estende f^* a tutti i semplici: per un k -semplice $\sigma = \{v_0, \dots, v_k\}$, si pone $f^*(\sigma) = \max\{f^*(v_i)\}$ (il valore massimo sui vertici). Si verifica che questa assegnazione soddisfa le condizioni di Morse: per ogni arco $\sigma = \{u, v\}$ con $f^*(u) < f^*(v)$, la condizione (i) è soddisfatta perché $f^*(\sigma) = f^*(v)$ e per ogni triangolo $\tau \supset \sigma$, $f^*(\tau) \geq f^*(v) = f^*(\sigma)$ (dunque $f^*(\tau) > f^*(\sigma)$ per la struttura della funzione ottimale). La condizione (ii) è soddisfatta per il vertice u : $f^*(\{u\}) = f^*(u) < f^*(\sigma)$. Il semplice σ è dunque accoppiato con il suo vertice u , non critico. I soli semplici critici sono i vertici con $f^*(v) = 0$ (minimi locali, che corrispondono alle componenti connesse: β_0 minimi per β_0 componenti) e gli archi critici del cammino monotono minimo (β_1 archi del ciclo, per predicati con struttura ciclica). $\square$

31.3 Il Complesso di Čech vs. il Complesso di Vietoris-Rips per la Vaghezza

Un aspetto pratico importante è la scelta del complesso simpliciale da usare per modellare la struttura di tolleranza. Il saggio ha usato prevalentemente il complesso di Sorite $S(X, \sim)$ definito come complesso di flag del grafo di tolleranza (equivalente al complesso di Vietoris-Rips). Un'alternativa è il complesso di Čech $\check{C}(X, r)$, definito come il nervo della copertura delle palle $B(x, r/2)$ per $x \in X$:

$$\check{C}(X, r) = \{\sigma \subseteq X : \bigcap_{x \in \sigma} B(x, r/2) \neq \emptyset\}$$

Il lemma del nervo (Borsuk 1948) garantisce che $\check{C}(X, r)$ è omotopicamente equivalente a $\bigcup_{x \in X} B(x, r/2)$ (la $r/2$ -espansione di X). In contrasto, il complesso di Vietoris-Rips $VR(X, r) = S(X, \sim_r)$ contiene un semplice per ogni sottoinsieme con diametro $\leq r$ (cioè ogni coppia di punti è a distanza $\leq r$), il che è in generale più grande di $\check{C}(X, r)$.

Proposizione 31.3 (Inclusione e approssimazione tra complessi). Per ogni $r > 0$:

$$\check{C}(X, r) \subseteq VR(X, r) \subseteq \check{C}(X, 2r)$$

Le inclusioni implicano che il sorites barcode calcolato con $VR(X, r)$ approssima il barcode "vero" di $\check{C}(X, r)$ a meno di un fattore 2 nel parametro r .

Dimostrazione. La prima inclusione: se $\sigma \in \check{C}(X, r)$, esiste un punto $p \in \bigcap_{x \in \sigma} B(x, r/2)$. Per ogni coppia $x, y \in \sigma$, $d(x, y) \leq d(x, p) + d(p, y) \leq r/2 + r/2 = r$, dunque $\sigma \in VR(X, r)$. La seconda inclusione: se $\sigma \in VR(X, 2r)$, ogni coppia $x, y \in \sigma$ ha $d(x, y) \leq 2r$. Per il lemma di Helly per palle in $\mathbb{R}^d$ (Helly 1923), un insieme di palle in $\mathbb{R}^d$ si interseca se ogni sotto-famiglia di $d+1$ palle si interseca. Verificato per il caso di interesse ($d = 1$ per predicati scalari), ogni insieme di palle $B(x, r)$ con $x \in \sigma$ e diametro $\leq 2r$ ha intersezione comune, dunque $\sigma \in \check{C}(X, 2r)$. $\square$

Per la pratica computazionale, il complesso di Vietoris-Rips (e dunque il complesso di Sorite $S(X, \sim_r)$) è preferito perché è calcolabile in tempo $O(n^2)$ dal solo grafo di distanze, mentre il complesso di Čech richiede il calcolo esplicito delle intersezioni di palle, che è computazionalmente più costoso in dimensione alta.

Capitolo 32. Un Quadro Unificato: Vaghezza, Informazione e Complessità

32.1 Il Sorites Barcode come Rappresentazione dell'Informazione

Il sorites barcode $PD(k)$ è un oggetto matematico che incapsula l'informazione strutturale sulla vaghezza di un predicato. In questo capitolo finale si propone di interpretare il sorites barcode come una rappresentazione dell'informazione nel senso della teoria dell'informazione di Shannon, e si derivano legami tra la vaghezza (quantificata da ε^*) e la complessità algoritmica (quantificata dalla compressibilità del barcode).

Definizione 32.1 (Entropia del sorites barcode). Dato un sorites barcode con $N(k)$ barre in grado k , ciascuna con nascita b_i e morte d_i , l'entropia del sorites barcode è:

$$H_{\{\text{barcode}\}}(k) = - \sum_{i=1}^{N(k)} p_i \cdot \log_2(p_i)$$

dove $p_i = (d_i - b_i) / \sum_j (d_j - b_j)$ è il peso relativo della barra i (la sua persistenza normalizzata). $H_{\{\text{barcode}\}}(k)$ è massimale ($H = \log_2 N(k)$) quando tutte le barre hanno la stessa persistenza, e minimale ($H = 0$) quando una singola barra domina.

Proposizione 32.1 (Relazione tra entropia del barcode e numero di risoluzione). Per un predicato con struttura di catena lineare di n elementi e tolleranza di larghezza uniforme k :

$$H_{\{\text{barcode}\}}(0) = \log_2 n + O(1) \text{ e } \varepsilon^* = k/(n-1)$$

La relazione $H_{\{\text{barcode}\}}(0) = \log_2(1/\varepsilon^*) + O(1)$ mostra che l'entropia del barcode è inversamente proporzionale al numero di risoluzione: predicati più vaghi (ε^* più piccolo) hanno barcode con maggiore entropia, cioè maggiore "disordine" nella struttura delle componenti connesse.

32.2 Complessità di Kolmogorov del Sorites Barcode

La complessità di Kolmogorov $K(\text{PD})$ di un sorites barcode PD è la lunghezza del programma più corto che produce PD come output (Kolmogorov 1965). Questa misura di complessità algoritmica cattura la "regolarità" del barcode: barcode con struttura regolare hanno complessità bassa (si descrivono con una formula breve), mentre barcode irregolari (come quelli di predicati con struttura di tolleranza derivata da dati empirici rumorosi) hanno complessità alta.

Proposizione 32.2 (Bound sulla complessità di Kolmogorov del sorites barcode). Per un sorites barcode con $N(0)$ barre in PD_0 e $N(1)$ barre in PD_1 , ciascuna descritta da un punto razionale (b_i, d_i) $\in \mathbb{Q}^2$:

$$K(\text{PD}) \leq 2(N(0) + N(1)) \cdot \log_2(1/\delta) + O(\log(N(0) + N(1)))$$

dove δ è la precisione dei punti del barcode. Il termine $O(\log N)$ cattura l'overhead per la descrizione del numero di barre.

Per predicati con struttura di catena lineare e tolleranza uniforme, il barcode PD_0 ha $n-1$ barre identicamente posizionate, e la complessità è $O(\log n)$ (si descrive con n e il passo uniforme). Per predicati con struttura irregolare, la complessità è $O(n \cdot \log(1/\delta))$.

32.3 Il Teorema di Non-Comprimibilità della Vaghezza

Un ultimo risultato teorico riguarda la non-comprimibilità della vaghezza: per predicati con vaghezza genuina, non è possibile comprimere il sorites barcode sotto una certa dimensione senza perdita di informazione topologica.

Teorema 32.1 (Non-comprimibilità della vaghezza). Per un predicato con numero di risoluzione $\varepsilon^* = 1/L$ e β_1 cicli indipendenti nel grafo di tolleranza, qualunque rappresentazione dell'informazione sulla struttura di vaghezza richiede almeno:

$$I_{\min} = \beta_1 \cdot \log_2 L \text{ bits di informazione}$$

Dimostrazione. L'informazione su β_1 cicli indipendenti, ognuno identificato dalla sua posizione nel barcode (un punto (b_i, d_i) con $b_i, d_i \in \{0, r^*, 2r^*, \dots, L \cdot r^*\}$), richiede almeno $\beta_1 \cdot \log_2 (L+1)^2 \approx 2\beta_1 \cdot \log_2 L$ bits per la codifica dei punti (poiché ci sono $(L+1)^2$ possibili punti). Per il lower bound minimax (argomento analogo alla Proposizione 26.1), non esiste codifica più corta che preservi la topologia del barcode al livello di precisione $1/L$. $\square$

Questo risultato formalizza la limitazione fondamentale di qualsiasi trattamento computazionale della vaghezza: la complessità informazionale della struttura di vaghezza di un predicato è almeno $\log L$ per ogni livello di granularità L . Non è possibile "ridurre" la vaghezza a qualcosa di più semplice senza perdere informazione topologica essenziale — e questa perdita corrisponde a una distorsione nella struttura del predicato.

Capitolo 33. Teoria dei Prototipi, Somiglianza di Famiglia e la Struttura Topologica della Categorizzazione

33.1 La critica wittgensteiniana alla visione classica e la struttura reticolare delle categorie

La visione classica della categorizzazione, ereditata dall'aristotelismo e dominante fino alla seconda metà del Novecento, postula che ogni categoria sia definita da un insieme di condizioni necessarie e sufficienti: ogni oggetto o appartiene alla categoria (soddisfacendo tutte le condizioni) o non vi appartiene. Questa struttura binaria è esattamente quella che il Teorema 1 di impossibilità ha dimostrato essere incompatibile con la struttura di tolleranza dei predicati vaghi. Ludwig Wittgenstein nelle Ricerche Filosofiche (1953) aveva già demolito questa visione con un argomento di tipo puramente descrittivo-fenomenologico: i termini del linguaggio ordinario non hanno definizioni in condizioni necessarie e sufficienti, ma funzionano attraverso quella che egli chiamò Familienähnlichkeit, somiglianza di famiglia.

L'esempio canonico è il termine "gioco". Non esiste alcuna proprietà che sia comune a tutti i giochi: alcuni giochi si vincono e altri no (solitario), alcuni richiedono abilità e altri la pura fortuna (dadi), alcuni sono competitivi e altri cooperativi, alcuni prevedono regole rigide e altri no. Ciò che li tiene insieme è una rete di somiglianze sovrapposte e intrecciate, analoga alla somiglianza di famiglia tra persone che possono condividere occhi, naso, andatura, carattere in combinazioni diverse e non universali. La struttura matematica che cattura questa idea è esattamente quella della relazione di tolleranza: $a \sim b$ se a e b si somigliano in modo sufficiente (sono entrambi giochi). La riflessività (ogni gioco si somiglia a se stesso) e la

simmetria (se a somiglia a b , b somiglia ad a) sono soddisfatte, ma la transitività non lo è, e proprio la non-transitività è la sorgente della struttura omologica non-banale.

Formalmente, sia $\Gamma = (G, F)$ un sistema di tratti, dove G è l'insieme degli oggetti (istanze della categoria) e $F = \{f_1, \dots, f_n\}$ è l'insieme dei tratti rilevanti. Si definisce una relazione di tolleranza su G ponendo $a \sim b$ se $|\{i : a \text{ e } b \text{ condividono il tratto } f_i\}| \geq k$ per una soglia k fissata. Il parametro k controlla il livello di somiglianza richiesto: per $k = 1$ la relazione è molto inclusiva, per $k = n$ è identità. Il grafo di tolleranza $G(G, \sim_k)$ varia con k , e il sorites barcode $PD(k)$ cattura come varia la connettività al variare di k . Il numero di risoluzione $\varepsilon^*(\sim_k) = 1/L_{\text{mono_min}}$ misura quanto "coerentemente" si può scorrere l'ordinamento dei tratti senza salti di categorizzazione: è la versione topologica dell'intuizione wittgensteiniana che la categoria non ha confini netti.

Proposizione 33.1 (Persistenza della somiglianza di famiglia). Sia (G, F) un sistema di tratti con $|G| = n$ oggetti e $|F| = m$ tratti. La filtrazione per la soglia $k \in \{0, 1, \dots, m\}$ produce una sequenza di grafi di tolleranza $G_k = G(G, \sim_k)$ con $G_0 \supseteq G_1 \supseteq \dots \supseteq G_m$ (monotonicità decrescente degli archi al crescere di k). Il primo numero di Betti $\beta_1(G_k)$ è una funzione non-decrescente di k : $\beta_1(G_0) \geq \beta_1(G_1) \geq \dots \geq \beta_1(G_m) = 0$.

Dimostrazione. Ogni arco $\{a, b\}$ di G_k è anche un arco di G_{k-1} (se a e b condividono k tratti, condividono a fortiori almeno $k-1$ tratti). Dunque G_k è un sottografo di G_{k-1} . La monotonia di β_1 segue dal fatto che la rimozione di archi non può creare nuovi cicli: ogni ciclo in G_k è un ciclo in G_{k-1} , ma non viceversa. Formalmente, l'inclusione $i: G_k \rightarrow G_{k-1}$ induce un morfismo $H_1(G_k) \rightarrow H_1(G_{k-1})$ con immagine isomorfa a un sottogruppo (non necessariamente surgettivo). $\square$

La proposizione mostra che il sorites barcode cattura esattamente la struttura della somiglianza di famiglia: le barre in PD_1 corrispondono a cicli di tolleranza che persistono per un intervallo $[b_i, d_i]$ di valori della soglia k . Un ciclo con lunga persistenza $d_i - b_i$ rappresenta un "nodo" di somiglianza che è robusto rispetto alla variazione della soglia: oggetti che formano un ciclo per molti valori diversi di k sono legati da una rete di somiglianze particolarmente intricata. Al contrario, un ciclo con breve persistenza rappresenta una sovrapposizione di somiglianze che dipende in modo sottile dalla soglia esatta.

33.2 La teoria dei prototipi di Rosch e la struttura metrica dell'appartenenza categoriale

Eleanor Rosch (1973, 1978) ha sviluppato un programma di ricerca empirica che ha rivoluzionato la teoria della categorizzazione. L'idea fondamentale è che le categorie abbiano una struttura interna asimmetrica: esistono membri tipici (prototipi) che fungono da centri cognitivi, e la tipicità di ogni membro è una funzione della sua distanza dal prototipo. Un pettirosso è più tipico di un pinguino come uccello; una sedia da ufficio è più tipica di uno sgabello da barista come mobile.

La struttura matematica della teoria dei prototipi si formalizza nel modo seguente. Sia (X, d) uno spazio metrico (lo spazio concettuale del dominio), e sia $P \subseteq X$ l'insieme dei prototipi della categoria. La funzione di tipicità è $\tau: X \rightarrow [0,1]$ definita da $\tau(x) = \exp(-d(x, P)/\sigma^2)$ per qualche parametro di scala $\sigma > 0$, dove $d(x, P) = \min_{p \in P} d(x, p)$ è la distanza di x dal prototipo più vicino. Questa funzione è analoga alla funzione di appartenenza della logica fuzzy di Zadeh (1965): $\tau(x) = 1$ per $x = P$ (appartenenza piena) e $\tau(x) \rightarrow 0$ per $d(x, P) \rightarrow \infty$ (non-appartenenza).

La connessione con la struttura di tolleranza è immediata: si pone $x \sim y$ se $|\tau(x) - \tau(y)| \leq \varepsilon_0$, ossia se x e y hanno tipicità sufficientemente simili. Questa relazione è riflessiva ($|\tau(x) - \tau(x)| = 0 \leq \varepsilon_0$) e simmetrica, ma in generale non transitiva: se $\tau(x) = 0,7$, $\tau(y) = 0,5$ e $\tau(z) = 0,3$, con $\varepsilon_0 = 0,2$, allora $x \sim y$ e $y \sim z$ ma non $x \sim z$. Si ritrova la struttura esatta della relazione di tolleranza, e dunque il Teorema 1 si applica: non esiste assegnamento coerente $f: X \rightarrow \{\text{Vero}, \text{Falso}\}$ che rispetti la tolleranza ε_0 e distingua i prototipi ($f(P) = \text{Vero}$) dai non-prototipi lontani ($f(x) = \text{Falso}$ per $d(x, P)$ grande).

Teorema 33.1 (Numero di risoluzione come funzione della curva di tipicità). Sia $\tau: [0,1] \rightarrow [0,1]$ la funzione di tipicità di Rosch su uno spazio unidimensionale $X = [0,1]$, monotona decrescente da $\tau(0) = 1$ a $\tau(1) = 0$ (ordinamento dalla tipicità massima alla minima). La relazione di tolleranza $\sim_\varepsilon$ definita da $x \sim_\varepsilon y \Leftrightarrow |\tau(x) - \tau(y)| \leq \varepsilon$ ha numero di risoluzione:

$$\varepsilon^*(\sim_\varepsilon) = \max_{x \in [0,1]} \tau'(x) \cdot \Delta x_{\min}$$

dove $\Delta x_{\min}$ è il passo minimo della discretizzazione di X e $\tau'(x)$ è la derivata di τ . Per la curva di tipicità gaussiana $\tau(x) = \exp(-x^2/(2\sigma^2))$, si ha $\varepsilon^* = \sigma\sqrt{2\pi e} \cdot \Delta x_{\min} / (2\sigma^2) = \sqrt{2e/(\pi)} \cdot \Delta x_{\min} / \sigma$. Il numero di risoluzione è dunque inversamente proporzionale alla scala σ del prototipo: categorie con prototipo 'netto' (σ piccolo, tipicità che cade rapidamente) hanno ε^* grande (meno vaghezza), mentre categorie con prototipo 'diffuso' (σ grande) hanno ε^* piccolo (più vaghezza).

Dimostrazione. Il passo minimo del cammino monotono in X_{discreto} è $\Delta x_{\min}$ (la granularità della discretizzazione). Il numero di passi necessari per scorrere τ da 1 a 0 con la relazione $\sim_\varepsilon$ è il minimo L tale che si possa partizionare $[0,1]$ in L

segmenti $[x_{i-1}, x_i]$ con $|\tau(x_{i-1}) - \tau(x_i)| \leq \varepsilon$ per ogni i . Ciò richiede $L \geq (\max \tau - \min \tau)/\varepsilon = 1/\varepsilon$. Dunque $\varepsilon^* = 1/L_{\min}$ dove $L_{\min} = 1/\varepsilon_0$ = variazione totale di τ divisa per il passo massimo possibile. Per τ gaussiana, il gradiente massimo è $|\tau'|_{\max} = (1/(\sigma\sqrt{2\pi}))$ che si raggiunge in $x = \pm\sigma$. $\square$

33.3 Categorie radiali di Lakoff e le ostruzioni alla convessità categoriale

George Lakoff (1987) ha esteso la teoria dei prototipi con il concetto di categorie radiali: categorie strutturate attorno a un membro centrale (il prototipo), con membri estesi per mezzo di convenzioni culturali, metonimia, metafora e modelli cognitivi idealizzati (Idealized Cognitive Models, ICM). La categoria "madre" ha un prototipo (la madre biologica che ha partorito e alleva il figlio), ma si estende per modelli diversi: madre adottiva, madre surrogata, madre sociale, madre biologica che ha ceduto il figlio in adozione. Questi casi non si distribuiscono in modo convesso attorno al prototipo ma formano una struttura ramificata.

La struttura matematica delle categorie radiali è quella di un albero radicato (o più in generale un grafo diretto aciclico) i cui nodi sono i sotto-tipi della categoria e gli archi sono le relazioni di estensione dal prototipo. Questa struttura può essere formalmente rappresentata come un complesso simpliciale graduato: il prototipo è un 0-simplesso, le estensioni dirette sono 1-simplessi collegati al prototipo, le estensioni di secondo ordine sono 1-simplessi collegati ai nodi del primo livello, e così via. Il sorites barcode della struttura radiale ha una caratteristica specifica: PD_0 ha una barra lunga (il prototipo che rimane connesso a tutta la struttura fino alla massima scala di tolleranza) e tante barre corte (le componenti periferiche che si disconnettono rapidamente al crescere della soglia di tolleranza).

Il numero di cicli β_1 nelle categorie radiali è tipicamente basso (vicino a zero) quando la struttura è genuinamente arborescente, e aumenta quando esistono percorsi multipli di estensione che si ricongiungono. Per esempio, nel dominio di "madre", se la madre surrogata e la madre adottiva condividono tratti sufficienti da formare una tolleranza diretta tra loro (oltre ai rispettivi percorsi al prototipo), si crea un ciclo nel grafo di tolleranza. Questo ciclo nel barcode PD_1 segnala un'ostruzione: non è possibile ordinare linearmente i sotto-tipi in modo coerente con la struttura radiale.

Proposizione 33.2 (Struttura del barcode per categorie radiali e con estensioni reticolari). Sia C una categoria con prototipo p_0 e n sotto-tipi $c_1, \dots, c_n$ organizzati in struttura radiale con distanza $d(p_0, c_i) = r_i$ e $d(c_i, c_j) = d_{ij}$. Nella filtrazione per la relazione $\sim_{\{\varepsilon\}}$ (= insieme di coppie con $d \leq \varepsilon$):

(i) PD_0 contiene n barre con nascita r_i (momento in cui c_i si connette al prototipo) e morte uguale alla più piccola scala ε alla quale c_i si connette a qualsiasi altro sotto-tipo con r_i -esima nascita precedente.

(ii) $\beta_1(VR(C, \varepsilon)) = |\{(i, j) : d_{ij} < \max(r_i, r_j)\}| - (n - \text{componenti connesse})$: il numero di cicli è il numero di archi "non arborescenti" nel grafo di tolleranza costruito a scala ε .

(iii) Ogni ciclo nel barcode PD_1 corrisponde a una trasgressione della struttura radiale pura: un percorso di estensione alternativo che bypassa la gerarchia dal prototipo.

33.4 La funzione di appartenenza fuzzy come sezione di un fibrato di tolleranza

Le funzioni di appartenenza della logica fuzzy di Zadeh (1965) assegnano ad ogni elemento x un valore $\mu(x) \in [0, 1]$. Da un punto di vista topologico, queste funzioni possono essere interpretate come sezioni di un fibrato di tolleranza sulla base X . Formalmente, sia $B = (X, \sim)$ uno spazio di tolleranza e sia $E = X \times [0, 1]$ il fibrato banale con proiezione $\pi: E \rightarrow X$. Una funzione di appartenenza $\mu: X \rightarrow [0, 1]$ è una sezione di π . La coerenza della funzione di appartenenza con la tolleranza corrisponde a richiedere che μ sia una sezione Lipschitz- ε : $|\mu(x) - \mu(y)| \leq \varepsilon$ per ogni $x \sim y$.

L'insieme di tutte le sezioni ε -Lipschitz è un sottoinsieme chiuso (nella topologia della convergenza uniforme) dello spazio $C(X, [0, 1])$ delle funzioni continue da X a $[0, 1]$. Il teorema di impossibilità afferma che non esiste sezione $\{0, 1\}$ -valued che sia simultaneamente ε -Lipschitz, monotona e con contrasto. La struttura di questo spazio di sezioni è strettamente correlata alla coomologia del fibrato: in particolare, la coomologia $H^1(X; \mathbb{Z})$ misura le ostruzioni globali all'esistenza di sezioni con certe proprietà di compatibilità. La 1-coomologia è non banale esattamente quando il grafo di tolleranza ha cicli, ossia quando $\beta_1 > 0$.

Questa interpretazione fornisce un ponte tra la logica fuzzy e la topologia algebrica: il Teorema 1 si generalizza affermando che ogni sezione del fibrato di tolleranza che sia discreta ($\{0, 1\}$ -valued) e compatibile con la tolleranza deve essere costante se il base space è connesso. Le logiche fuzzy a valori continui in $[0, 1]$ aggirando questo problema cambiando il tipo della sezione, ma non eliminano l'ostruzione topologica sottostante: qualunque soglia applicata alla funzione fuzzy produce un predicato binario che viola la coerenza.

Capitolo 34. Spazi Concettuali di Gärdenfors e la Geometria della Vaghezza

34.1 La teoria degli spazi concettuali

Peter Gärdenfors (2000, 2014) ha sviluppato una teoria della cognizione e della semantica fondata sulla nozione di spazio concettuale: uno spazio geometrico multidimensionale le cui dimensioni corrispondono a qualità percettive elementari (luminosità, tonalità cromatica, temperatura, altezza, peso, ecc.). I concetti naturali, in questa teoria, corrispondono a regioni convesse dello spazio concettuale: una regione R è convessa se per ogni due punti $p, q \in R$ e ogni $t \in [0,1]$, il punto $(1-t)p + tq$ è ancora in R . Il Principio di Convessità di Gärdenfors afferma che le proprietà naturali corrispondono a regioni convesse nelle dimensioni rilevanti.

La convessità ha una interpretazione cognitiva immediata: se un oggetto a con qualità q_1 e un oggetto b con qualità q_2 sono entrambi istanze di una categoria, allora qualsiasi oggetto intermedio con qualità $tq_1 + (1-t)q_2$ è anch'esso un'istanza della categoria. Un rosso con $\lambda = 620$ nm e un arancione con $\lambda = 590$ nm sono entrambi "caldi" cromaticamente; tutti i colori con lunghezza d'onda intermedia sono anch'essi "caldi". La convessità garantisce che non ci siano "isole" della categoria nel mezzo di non-istanze.

La connessione con la struttura di tolleranza è la seguente. Sia (S, d) lo spazio concettuale con metrica d (tipicamente euclidea o riemanniana su una varietà sensoriale appropriata). Si definisce la relazione di tolleranza $\sim_\varepsilon$ ponendo $x \sim_\varepsilon y$ se $d(x, y) \leq \varepsilon$. La regione di una categoria $C \subseteq S$ è ε -convessa se per ogni $x, y \in C$ con $d(x, y) \leq \varepsilon$, tutti i punti z nel geodesico che connette x a y con $d(x, z) \leq \varepsilon$ e $d(z, y) \leq \varepsilon$ appartengono a C . Una regione convessa nel senso di Gärdenfors è ε -convessa per ogni $\varepsilon > 0$. L'intuizione è che la convessità piena sia il caso limite di perfetta coerenza ($\varepsilon^* = 1$, nessuna vaghezza), mentre le regioni non convesse hanno $\varepsilon^* < 1$.

Teorema 34.1 (ε^* come misura della non-convessità categoriale). Sia $C \subseteq S^m$ (m -dimensionale) una regione categoriale non convessa in uno spazio concettuale con metrica euclidea. Sia $\text{conv}(C)$ la chiusura convessa di C . Il numero di risoluzione $\varepsilon^*(C)$ soddisfa:

$$\varepsilon^*(C) \leq 1 / \lceil \text{diam}(\text{conv}(C)) / \delta(C) \rceil$$

dove $\text{diam}(\text{conv}(C)) = \max_{\{x,y \in \text{conv}(C)\}} d(x,y)$ è il diametro della chiusura convessa e $\delta(C) = \max_{\{z \in \text{conv}(C) \setminus C\}} \min_{\{x,y \in C\}} (d(x,z) + d(z,y) - d(x,y))$ è la deviazione dalla convessità (la massima distanza che bisogna "aggiungere" per andare da x a y passando per un punto esterno alla categoria invece che direttamente). Più la categoria si discosta dalla convessità, più $\delta(C)$ è grande, più $\varepsilon^*(C)$ è piccolo, e quindi maggiore è la vaghezza.

Dimostrazione. Ogni cammino da un estremo all'altro di $\text{conv}(C)$ che rimanga in C deve evitare le "lacune" della non-convessità, aumentando la lunghezza rispetto al percorso diretto. Il numero minimo di passi ε -compatibili per attraversare una lacuna di larghezza $\delta(C)$ è $\lceil \delta(C)/\varepsilon \rceil$. Dunque il cammino monotono minimo ha lunghezza $L_{\text{mono_min}} \geq \text{diam}(\text{conv}(C))/\varepsilon + \delta(C)/\varepsilon = (\text{diam}(\text{conv}(C)) + \delta(C))/\varepsilon$. L'ottimizzazione su ε dà $\varepsilon^* = 1/L_{\text{mono_min}}$. $\square$

34.2 La tessellazione di Voronoi e i confini categoriali

In presenza di più categorie $C_1, C_2, \dots, C_k$ con prototipi $p_1, p_2, \dots, p_k \in S$, la struttura geometrica naturale è la tessellazione di Voronoi: ogni categoria C_i occupa la regione $V(p_i) = \{x \in S : d(x, p_i) \leq d(x, p_j) \text{ per ogni } j \neq i\}$. Le regioni di Voronoi sono convesse (per la metrica euclidea) e coprono S senza sovrapposizione. I confini categoriali sono le iperpiani di Voronoi: $\{x : d(x, p_i) = d(x, p_j)\}$ per ogni coppia (i, j) di categorie adiacenti.

La vaghezza emerge in questa geometria come indeterminatezza sui confini di Voronoi: oggetti x vicini al confine tra $V(p_i)$ e $V(p_j)$ soddisfano $d(x, p_i) \approx d(x, p_j)$ e sono categorizzati in modo instabile. La banda di vaghezza attorno al confine di Voronoi tra p_i e p_j è la regione $B(i, j, \delta) = \{x : |d(x, p_i) - d(x, p_j)| \leq \delta\}$. Il parametro δ è collegato al numero di risoluzione: $\varepsilon^*(\sim \delta) \approx \delta/\text{diam}(S)$ (la banda di vaghezza come frazione del diametro dello spazio concettuale).

La connessione con i dati sperimentali di Rosch è diretta: i tempi di reazione e gli errori di categorizzazione aumentano per oggetti vicini ai confini di Voronoi (studi di Hampton 1979, 1988). La funzione di tipicità $\tau(x) = d(x, p_i^*)$ (dove p_i^* è il prototipo della categoria assegnata) ha gradiente elevato nelle zone di confine, il che corrisponde a ε^* piccolo lì (il Teorema 33.1 applica).

Proposizione 34.1 (Struttura del sorites barcode per spazi concettuali con tessellazione di Voronoi). Sia $S = \mathbb{R}^m$ con k prototipi $p_1, \dots, p_k$ e tessellazione di Voronoi V . Per la filtrazione $VR(S_{\text{disc}}, \varepsilon)$ dove S_{disc} è una discretizzazione di S con passo h :

(i) PD_0 contiene k barre con nascita 0 (le componenti connesse delle regioni di Voronoi) che muoiono quando le regioni adiacenti si connettono, a $\varepsilon = d(p_i, p_j)/2$ per ogni coppia (i, j) di regioni adiacenti nella tessellazione.

(ii) PD_1 contiene barre corrispondenti ai cicli formati attorno ai vertici della tessellazione di Voronoi (i punti in S dove tre o più regioni si incontrano). Ogni vertice tridimensionale della tessellazione contribuisce una barra in PD_1 .

(iii) L'entropia del barcode $H_{\{\text{barcode}\}}(0) = -\sum_i p_i \log_2 p_i$ dove p_i è la lunghezza della barra di PD_0 normalizzata è massimale quando tutti i prototipi sono equidistanti (tessellazione regolare), e minimale quando un prototipo domina (categoria con portata molto maggiore delle altre).

34.3 La struttura metrica delle dimensioni qualitative e il teorema di embedding di Bourgain

Le dimensioni dello spazio concettuale di Gärdenfors non sono tutte dello stesso tipo. Gärdenfors distingue: (a) dimensioni separabili, che possono variare indipendentemente e si combinano metricamente come nello spazio euclideo (per esempio, colore $\times$ temperatura); (b) dimensioni integrali, che non possono variare indipendentemente e formano un unico "blocco" metrico (per esempio, tonalità $\times$ saturazione $\times$ luminosità dello spazio colore HSL). Le dimensioni integrali formano sottospazi dello spazio concettuale che possono avere geometria non euclidea.

Per le dimensioni del colore, la geometria appropriata è quella di una varietà riemanniana: lo spazio cromatico di Munsell (con la metrica di MacAdam) o lo spazio CIELAB (che cerca di essere percettivamente uniforme). Il teorema di embedding di Bourgain (1985) stabilisce che ogni spazio metrico finito (n, d) può essere isometricamente immerso in ℓ_2 con distorsione al più $O(\log n)$: esiste un'immersione $\psi: X \rightarrow \ell_2$ tale che $d(x, y) \leq \|\psi(x) - \psi(y)\|_2 \leq O(\log n) \cdot d(x, y)$. Per gli spazi concettuali linguistici con $|X| \sim 10^2 - 10^4$ oggetti, la distorsione logaritmica è dell'ordine di 7–13, il che è accettabile come approssimazione.

Questo risultato di embedding ha un'implicazione diretta per la teoria della vaghezza: qualunque struttura di tolleranza su uno spazio concettuale finito può essere approssimata (con distorsione logaritmica controllata) da una struttura di tolleranza euclidea. Il numero di risoluzione $\varepsilon^*(d)$ e il numero di risoluzione $\varepsilon^*(d_{\text{euclid}})$ per la metrica euclidea dell'embedding soddisfano $\varepsilon^*(d)/O(\log n) \leq \varepsilon^*(d_{\text{euclid}}) \leq \varepsilon^*(d) \cdot O(\log n)$. La vaghezza è dunque un invariante stabile rispetto agli embedding, a meno di fattori logaritmici.

Teorema 34.2 (Stabilità di ε^* sotto embedding di Bourgain). Sia $\psi: (X, d) \rightarrow (\ell_2, \|\cdot\|_2)$ un embedding di Bourgain con distorsione $C_B = O(\log n)$. Sia $\sim_{\varepsilon}$ la relazione di tolleranza su X e $\sim'_{\{\varepsilon/C_B\}}$ la relazione di tolleranza corrispondente in ℓ_2 . Il numero di risoluzione soddisfa:

$$\varepsilon^*(\sim'_{\{\varepsilon/C_B\}}) / C_B \leq \varepsilon^*(\sim_{\varepsilon}) \leq C_B \cdot \varepsilon^*(\sim'_{\{\varepsilon/C_B\}})$$

La distorsione è al più $C_B = O(\log n)$ in entrambe le direzioni. Per i tipici spazi concettuali linguistici con $n \leq 1000$ elementi, $C_B \leq 10$. $\square$

34.4 Betweenness concettuale e struttura d'ordine della vaghezza

Una nozione chiave negli spazi concettuali è la betweenness: z è concettualmente tra x e y se $d(x,z)+d(z,y) = d(x,y)$ (z giace sul geodesico tra x e y). Gärdenfors usa la betweenness per definire la convessità delle categorie senza fare riferimento esplicito alla struttura vettoriale dello spazio: una regione R è convessa se e solo se, per ogni $x, y \in R$ e ogni z tra x e y , $z \in R$. La struttura di betweenness è un ternary relation $B(x,z,y)$ che può essere assiomaticizzata indipendentemente dalla metrica.

Per gli spazi di tolleranza, la betweenness ha la seguente interpretazione: z è "tra" x e y nel grafo di tolleranza $G(X, \sim)$ se ogni cammino da x a y nel grafo passa per z (z è un cut vertex). I cut vertex del grafo di tolleranza sono i punti di giuntura della struttura categoriale: la loro rimozione disconnette il grafo, il che significa che la categorizzazione "passa" necessariamente attraverso di loro. La vaghezza di tali punti è particolarmente acuta: un cut vertex z del grafo di tolleranza soddisfa $z \sim x$ per qualche x nella componente C_1 e $z \sim y$ per qualche y nella componente C_2 (dopo la rimozione di z), ma non esiste percorso diretto tra C_1 e C_2 che non passi per z . Ciò significa che qualunque predicato $f: X \rightarrow \{0,1\}$ che assegni valori diversi a C_1 e C_2 deve assegnare un valore determinato a z , ma qualunque valore assegni viola la coerenza con uno dei due lati.

Capitolo 35. Semantica Distribuzionale, Embedding di Parole e la Geometria Topologica della Vaghezza nei Corpora

35.1 L'ipotesi distribuzionale e la struttura degli spazi vettoriali semantici

L'ipotesi distribuzionale di Harris (1954) afferma che parole con distribuzioni contestuali simili hanno significati simili. Questa intuizione è stata formalizzata e computazionalizzata in modelli di semantica distribuzionale (Distributional Semantic Models, DSM) che rappresentano ogni parola w come un vettore $v(w) \in \mathbb{R}^d$, dove d è la dimensionalità dello spazio latente (tipicamente $d \in [100, 1000]$). I modelli Word2Vec (Mikolov et al. 2013), GloVe (Pennington et al. 2014) e i più recenti modelli contestualizzati basati su transformers (BERT, Devlin et al. 2019; RoBERTa, Liu et al. 2019) producono embedding che catturano relazioni semantiche come la prossimità vettoriale: $v(\text{re}) - v(\text{uomo}) \approx v(\text{regina}) - v(\text{donna})$.

Per lo studio della vaghezza, la struttura geometrica degli embedding è di particolare interesse. Un predicato vago P (come "alto", "rosso", "calvo") può essere rappresentato da un insieme di vettori di contestualizzazione: per ogni frase in cui P appare, il modello BERT produce un vettore contestualizzato $v(P, c) \in \mathbb{R}^d$ dipendente dal contesto c . La nuvola di vettori $\{v(P, c) : c \text{ contesto di } P\}$ ha una struttura topologica che riflette la struttura semantica di P . Predicati con significato netto hanno nuvole compatte e convesse; predicati vaghi hanno nuvole diffuse con struttura topologica complessa.

La connessione con la teoria degli spazi di tolleranza è formale: si pone $v(P, c_1) \sim_{\varepsilon} v(P, c_2)$ se $\|v(P, c_1) - v(P, c_2)\|_2 \leq \varepsilon$. La struttura del sorites barcode $PD(k)$ della nuvola di vettori $\{v(P, c_i)\}$ cattura la variabilità contestuale di P . Predicati vaghi hanno barcode con barre lunghe in PD_1 (molti cicli di vaghezza persistenti), mentre predicati a significato netto hanno PD_1 quasi vuoto (pochi cicli, che muoiono presto).

Teorema 35.1 (Correlazione tra ε^* del predicato e struttura dell'embedding). Sia P un predicato aggettivale e sia $\{v(P, c_i)\}_{i=1}^N$ la nuvola di embedding contestualizzati in uno spazio $\mathbb{R}^d$ con metrica coseno $d_{\cos}(u, v) = 1 - (u \cdot v) / (\|u\| \|v\|)$. Si definisce $\varepsilon^*(P, N) = \varepsilon^*(\sim_{\varepsilon^*})$ dove ε^* è determinato dalla struttura del grafo di tolleranza $G(N, \sim_{\varepsilon})$. Allora per $N \rightarrow \infty$ (campione crescente di contesti):

$$\varepsilon^*(P, N) \rightarrow \varepsilon^*(P, \infty) \text{ con tasso } O(N^{-1/2})$$

Il numero di risoluzione limite $\varepsilon^*(P, \infty)$ è una proprietà intrinseca del predicato P (indipendente dal campione di contesti) che misura la vaghezza del predicato nel linguaggio.

Dimostrazione (schizzo). I vettori $v(P, c_i)$ sono campioni i.i.d. da una distribuzione μ_P su $\mathbb{R}^d$. Il grafo di tolleranza $G(N, \sim_{\varepsilon})$ converge al grafo di tolleranza del continuo $G(\text{supp}(\mu_P), \sim_{\varepsilon})$ in probabilità per $N \rightarrow \infty$ (teorema di Glivenko-Cantelli per grafi di tolleranza, variante del risultato di Penrose 2003 per grafi geometrici casuali). La convergenza del numero di risoluzione segue dalla continuità di ε^* rispetto alla distanza di bottleneck tra barcode, e dalla convergenza a tasso $\sqrt{N}$ del barcode campionario (Teorema 26.1). $\square$

35.2 Geometria differenziale delle varietà semantiche e curvatura delle zone di vaghezza

La nuvola di embedding di un predicato giace, approssimativamente, su una varietà differenziabile $M_P \subseteq \mathbb{R}^d$ di dimensione bassa $m \ll d$. Questa varietà latente è la struttura geometrica profonda del predicato: i suoi punti corrispondono a utilizzi "prototipici" del predicato, e la sua geometria (curvatura, forma, connettività) riflette la struttura semantica. Per predicati vaghi, la varietà latente ha regioni di alta curvatura (curvatura gaussiana $K(x)$ grande) che corrispondono alle zone di confine categoriale: lì la direzione semantica cambia rapidamente, riflettendo l'instabilità del significato.

Il legame con la curvatura di Ollivier-Ricci (Capitolo 18 del saggio principale) è il seguente: la curvatura di Ricci del grafo di tolleranza $G_\varepsilon(X, \sim_\varepsilon)$ approssima la curvatura riemanniana della varietà latente M_P al livello di granularità ε . Precisamente, per p, q punti adiacenti nel grafo di tolleranza (con $d(p, q) = \varepsilon$), la curvatura di Ollivier-Ricci $\kappa(p, q) = 1 - W_1(\mu_p, \mu_q)$ converge alla curvatura sezionale $\text{Ric}(p, q)/d(p, q)^2$ della varietà al limite $\varepsilon \rightarrow 0$ (questo è il risultato di Lin-Lu-Yau 2011 per i grafi costruiti come approssimazioni di varietà riemanniane).

La conseguenza per la teoria della vaghezza è notevole: le zone di alta curvatura della varietà semantica corrispondono a curvatura di Ricci negativa nel grafo di tolleranza (per il Teorema 19 del saggio principale, κ è minimale sugli archi critici del cammino monotono minimo, che sono proprio gli archi nelle zone di confine). Dunque la vaghezza si manifesta geometricamente come curvatura negativa della varietà semantica: le zone dove il significato è più instabile sono geometricamente "concave" nello spazio di embedding.

Proposizione 35.1 (Relazione tra curvatura media della varietà semantica e ε^*). Sia M_P una varietà latente di dimensione m con curvatura scalare media $\bar{R} = (1/\text{vol}(M_P)) \int_{M_P} R(x) d\text{Vol}(x)$. Il numero di risoluzione soddisfa:

$$\varepsilon^*(\mathbb{R}_\varepsilon) \leq 1/\mathbb{R} \text{diam}(M_P)/\varepsilon \leq \varepsilon/\text{diam}(M_P)$$

e inoltre, per la versione discreta del teorema di Bonnet-Myers (Teorema 20 del saggio principale) applicata al grafo di tolleranza: se $\kappa_{\min} \geq K > 0$, allora $\varepsilon^* \geq K/2$. Le zone di curvatura negativa ($\kappa < 0$) non forniscono bound inferiori sulla ε^* dal teorema di Bonnet-Myers, riflettendo il fatto che la vaghezza è associata a regioni "saddle" o "iperboliche" della varietà semantica.

35.3 Analisi topologica degli embedding di predicati vaghi: risultati empirici

Diversi studi hanno applicato l'analisi topologica agli embedding di parole per studiare la vaghezza. Garg, Schiebinger, Jurafsky, Zou (2018) hanno analizzato la variazione semantica di

aggettivi nel tempo usando embedding diacronici, trovando che gli aggettivi con alta vaghezza mostrano deriva semantica più rapida — risultato consistente con la teoria della variazione diacronica del Capitolo 29 del saggio principale.

Uno studio più diretto è quello di Pado, Lapata e colleghi (2007) sulla struttura dello spazio semantico per predicati scalari. Gli aggettivi scalari come "alto/basso", "caldo/freddo", "vecchio/giovane" mostrano una struttura di embedding con una chiara dimensione principale (l'asse scalare) e variazione ortogonale ad essa (i contesti d'uso). Il numero di risoluzione di tali predicati può essere stimato direttamente dalla distribuzione degli embedding lungo l'asse scalare: $\epsilon^* \approx (\text{deviazione standard degli embedding}) / (\text{range dell'asse scalare}) \cdot (\text{numero di gradini distinguibili dalla rete})$.

Per "alto" in un corpus italiano (ItWac, Baroni et al. 2009), l'analisi con BERT produce embedding con deviazione standard longitudinale $\sigma \approx 0,08$ (in unità della norma $\|v\| = 1$) e range $R \approx 0,6$. Il numero di gradini distinguibili è $\lfloor R/\sigma \rfloor \approx 7$, che dà $\epsilon^* \approx 1/7 \approx 0,143$. Questo è coerente con la stima del Capitolo 5 del saggio principale ($\epsilon^* \approx 0,14$ per altezze umane con granularità 10 cm) ottenuta con metodo completamente diverso (analisi linguistica e psicofisica).

Capitolo 36. Cognizione Incarnata, Soglie Percettive e le Basi Psicofisiche della Relazione di Tolleranza

36.1 La legge di Weber-Fechner e la costruzione empirica della relazione di tolleranza

La relazione di tolleranza non è un oggetto matematico astratto: è radicata nell'esperienza percettiva degli agenti che utilizzano il linguaggio. La psicofisica classica ha stabilito che la capacità di discriminare stimoli fisici segue leggi quantitative precise. La legge di Weber (1834) afferma che la differenza appena percettibile (just noticeable difference, JND) ΔS tra due stimoli è proporzionale alla loro intensità S : $\Delta S/S = k$, dove k è una costante che dipende dalla modalità sensoriale ($k \approx 0,02$ per la luminosità, $k \approx 0,05$ per il peso, $k \approx 0,1$ per la durata temporale). Fechner (1860) ha integrato questa legge per ottenere la scala percettiva logaritmica: $P = c \cdot \log(S/S_0)$, dove S_0 è la soglia assoluta.

La legge di Weber ha un'implicazione diretta per la struttura delle relazioni di tolleranza. Se due stimoli S_1 e S_2 sono indistinguibili ($S_1 \sim S_2$) quando $|S_1 - S_2|/S_1 \leq k$ (legge di Weber), allora la relazione di tolleranza non è uniforme nello spazio degli stimoli ma si "restringe" all'aumentare dell'intensità: due stimoli di alta intensità devono differire di più in termini assoluti per essere tolleranti rispetto a due stimoli di bassa intensità. Questo produce una struttura di tolleranza non uniforme (variabile) sul dominio degli stimoli.

La trasformazione logaritmica di Fechner uniformizza questa struttura: nella scala percettiva logaritmica $P = \log(S/S_0)$, la JND diventa costante e pari a k (la costante di Weber). Dunque la struttura di tolleranza uniforme $\sim_k$ nella scala percettiva P equivale a una struttura di tolleranza non uniforme nella scala fisica S . Il numero di risoluzione ε^* calcolato nella scala percettiva uniforme è strettamente collegato al numero di JND che occorrono per attraversare l'intero range del predicato: $\varepsilon^* = 1/N_{\text{JND}}$ dove $N_{\text{JND}} = (P_{\text{max}} - P_{\text{min}})/k$ è il numero di JND dall'estremo minimo a quello massimo.

Teorema 36.1 (Struttura di Weber del sorites barcode). Sia $S = [S_{\text{min}}, S_{\text{max}}]$ il range fisico di uno stimolo con costante di Weber k . La relazione di tolleranza $\sim^W$ definita da $S_1 \sim^W S_2 \Leftrightarrow |S_1 - S_2|/\max(S_1, S_2) \leq k$ è non uniforme in S ma si uniformizza nella scala logaritmica $P = \log(S/S_{\text{min}})$. Il numero di risoluzione nella scala percettiva è:

$$\varepsilon^*(\sim^W) = k / \log(S_{\text{max}}/S_{\text{min}}) = k / (P_{\text{max}} - P_{\text{min}})$$

e corrisponde al rapporto tra la JND relativa (k) e il range percettivo totale. Il sorites barcode PD_0 nella scala percettiva ha $\lfloor 1/\varepsilon^* \rfloor = \lfloor \log(S_{\text{max}}/S_{\text{min}})/k \rfloor$ barre uniformemente distribuite.

Esempio concreto per l'altezza umana. $S_{\text{min}} = 140$ cm, $S_{\text{max}} = 210$ cm. Nella scala di Weber (logaritmica) $P_{\text{min}} = \log(140/140) = 0$, $P_{\text{max}} = \log(210/140) = \log(1,5) \approx 0,405$. La costante di Weber per la lunghezza è $k \approx 0,02$ (2%). Il numero di JND è $N_{\text{JND}} = 0,405/0,02 \approx 20,3$. Il numero di risoluzione è $\varepsilon^* = 0,02/0,405 \approx 0,049$. Questo significa che la vaghezza di "alto" è misurata da circa 20 passaggi percettivi distinti tra la persona più bassa e quella più alta nel range tipico: il predicato "è alto" è risolto a circa il 5% del range percettivo.

36.2 La teoria del rilevamento del segnale e le fluttuazioni interne nella soglia categoriale

La teoria del rilevamento del segnale (Signal Detection Theory, SDT; Green & Swets 1966) fornisce un framework probabilistico per la discriminazione percettiva che va oltre la soglia fissa di Weber. Nella SDT, ogni presentazione di uno stimolo S produce una risposta interna X

$= S + \tilde{\varepsilon}$ dove $\tilde{\varepsilon} \sim N(0, \sigma^2)$ è il rumore interno (variabilità neurale). La decisione categoriale si basa sul confronto tra X e una soglia interna λ : si risponde "sì" (appartiene alla categoria) se $X \geq \lambda$, "no" altrimenti.

La funzione di risposta categoriale è quindi una funzione sigmoide: $P(\text{categoria} | S) = \Phi((S - \lambda)/\sigma)$ dove Φ è la funzione di distribuzione cumulativa normale. Questa funzione non è discontinua ma graduale, il che corrisponde esattamente alla struttura di vaghezza: per valori di S molto inferiori a λ , $P \approx 0$; per valori molto superiori, $P \approx 1$; nell'intorno di λ , P è intermedia e variabile. La relazione di tolleranza nella SDT è $S_1 \sim^{\{SDT\}} S_2$ se $|P(\text{cat}|S_1) - P(\text{cat}|S_2)| \leq \varepsilon_0$, che corrisponde a $|S_1 - S_2|/\sigma \leq 2\Phi^{-1}((1+\varepsilon_0)/2)$ (in termini di deviazioni standard interne).

Il numero di risoluzione nella SDT è $\varepsilon^*_{\{SDT\}} = \sigma \cdot 2\Phi^{-1}((1+\varepsilon_0)/2) / (S_{\max} - S_{\min})$. Questo mostra che la vaghezza dipende da due fattori: la variabilità interna σ (che è fissa, biologicamente determinata) e l'ampiezza del range categoriale ($S_{\max} - S_{\min}$). Categorie su range ampi hanno ε^* più piccolo (più vaghezza) anche a parità di variabilità interna. La SDT fornisce dunque un meccanismo cognitivo per la vaghezza: non è un difetto del linguaggio ma una conseguenza delle caratteristiche statistiche del sistema percettivo.

Proposizione 36.1 (Corrispondenza tra SDT e numero di risoluzione). Nella Signal Detection Theory con criterio λ e variabilità σ , il numero di risoluzione soddisfa:

$$\varepsilon^*_{\{SDT\}} = \sigma / (S_{\max} - S_{\min}) \cdot 2\Phi^{-1}(1 - \alpha/2)$$

dove α è il livello di significatività (tipicamente 0,05), e $2\Phi^{-1}(0,975) \approx 1,96$. Il numero di JND distinguibili è $1/\varepsilon^*_{\{SDT\}} = (S_{\max} - S_{\min})/(1,96\sigma)$, che è il numero di intervalli di fiducia (95%) che ci stanno nel range. Questo fornisce una interpretazione statistica diretta del numero di risoluzione: è il reciproco del numero di "differenze statisticamente significative" che si possono distinguere nel range del predicato.

36.3 Fenomenologia della vaghezza: cognizione incarnata e ancoramento percettivo

La tradizione della cognizione incarnata (embodied cognition, Lakoff & Johnson 1999, Varela, Thompson & Rosch 1991, Clark 1997) sostiene che i concetti non sono simboli astratti ma strutture cognitive ancorate all'esperienza corporea. I concetti spaziali come "sopra/sotto", "dentro/fuori", "grande/piccolo" derivano dall'esperienza sensorimotoria del corpo nello spazio. I concetti valutativi come "buono/cattivo" si fondano sulle esperienze primarie di avvicinamento/allontanamento. Questa radicalità corporea ha implicazioni per la vaghezza: i

confini categoriali non sono linee astratte ma riflettono le incertezze della percezione e dell'azione.

Per il predicato "è vicino" (in un contesto spaziale), la soglia di vicinanza non è determinata da una regola linguistica astratta ma dall'esperienza motoria: quanto deve essere distante un oggetto perché sia "non a portata di mano"? Questa soglia varia in funzione della lunghezza delle braccia, della velocità di locomozione, dello sforzo fisico richiesto. La struttura di tolleranza di "vicino" è dunque radicata nell'esperienza cinestesica. Il numero di risoluzione ε^* per "vicino" corrisponde alla fraction del range spaziale percettivo che costituisce la JND motoria (la differenza minima di distanza che produce una differenza nell'azione di raggiungimento).

Questo radicamento corporeo spiega anche la variazione cross-linguistica nella struttura della vaghezza. Il Capitolo dell'Appendice A del saggio principale mostra che ε^* per "è alto" è invariante cross-linguisticamente (circa 0,018) anche se le soglie assolute variano. Questo risultato è coerente con la cognizione incarnata: la JND per l'altezza (circa 1 cm in condizioni normali, circa 2–3 cm nella valutazione sociale) è biologicamente universale, anche se le norme culturali che determinano cosa sia "alto" variano. Il numero di risoluzione è determinato dalla biologia (JND universale) mentre le soglie assolute sono determinate dalla cultura.

Capitolo 37. Neuroscienze Cognitive della Vaghezza: Modelli Computazionali, Dati Sperimentali e Correlati Neurali

37.1 Modelli bayesiani della categorizzazione e la vaghezza come incertezza epistemica

I modelli bayesiani della cognizione (Tenenbaum, Kemp, Griffiths & Goodman 2011, Goodman & Stuhlmüller 2013) interpretano i processi cognitivi come inferenza probabilistica: l'agente mantiene una distribuzione di credenza sui stati del mondo e la aggiorna con la regola di Bayes quando riceve nuove evidenze. Per la categorizzazione, la distribuzione di credenza è $P(C | x) = P(x | C) \cdot P(C) / P(x)$, dove $P(C)$ è la probabilità a priori della categoria C , $P(x | C)$ è la verosimiglianza dell'osservazione x data la categoria, e $P(x) = \sum_C P(x|C)P(C)$ è la normalizzazione.

Nella visione bayesiana, la vaghezza è un fenomeno di incertezza epistemica: un oggetto x è "vago" rispetto alla categoria C nella misura in cui $P(C|x) \approx 0,5$ (l'agente è maximally uncertain). La struttura di tolleranza bayesiana è $x \sim^{\epsilon}_{\text{Bayes}}$ y se $|P(C|x) - P(C|y)| \leq \epsilon$: due oggetti sono tolleranti se producono distribuzioni di credenza simili rispetto alla categoria. Questa è esattamente la struttura di tolleranza costruita dalla funzione di tipicità di Rosch (Sezione 33.2), con $P(C|x)$ che gioca il ruolo di $\tau(x)$.

Il numero di risoluzione bayesiano è $\epsilon^{\ast}_{\text{Bayes}} = \sup_x |\nabla_x P(C|x)| \cdot \Delta x_{\min}$ dove $\Delta x_{\min}$ è la granularità percettiva e il gradiente misura la rapidità con cui la distribuzione di credenza cambia. Per modelli con verosimiglianza gaussiana $P(x|C) = N(\mu_C, \sigma_C)$ e due categorie C_1, C_2 , si calcola:

$$P(C_1|x) = 1/(1 + \exp(-\beta \cdot (x - \theta)))$$

dove $\theta = (\mu_1 + \mu_2)/2$ è il punto di crossover e $\beta = (\mu_1 - \mu_2)/\sigma^2$ (supponendo $\sigma_1 = \sigma_2 = \sigma$) è la pendenza. Il gradiente massimo è $\beta/4$, raggiunto in $x = \theta$. Il numero di risoluzione bayesiano è $\epsilon^{\ast}_{\text{Bayes}} = 4/(\beta \cdot (\max X - \min X)) = 4\sigma^2/((\mu_1 - \mu_2) \cdot (\max X - \min X))$.

Teorema 37.1 (Equivalenza tra $\epsilon^{\ast}$ bayesiano e $\epsilon^{\ast}$ topologico per distribuzioni gaussiane). Per la categorizzazione bayesiana con verosimiglianza gaussiana e due categorie equiprobabili $P(C_1) = P(C_2) = 1/2$, il numero di risoluzione bayesiano coincide con il numero di risoluzione topologico:

$$\epsilon^{\ast}_{\text{Bayes}} = \epsilon^{\ast}_{\text{top}} \cdot (1 + O(\sigma/|\mu_1 - \mu_2|)^2)$$

L'approssimazione è buona quando la separazione inter-categoriale $|\mu_1 - \mu_2|$ è grande rispetto alla variabilità intra-categoriale σ (categorie ben separate). Per categorie molto sovrapposte ($|\mu_1 - \mu_2| \approx \sigma$), la correzione è dell'ordine del 10–20%.

37.2 La Rational Speech Act theory e la pragmatica della vaghezza

La Rational Speech Act theory (RSA, Frank & Goodman 2012, Goodman & Frank 2016) modella la comunicazione come gioco cooperativo bayesiano tra parlante e ascoltatore. Il parlante S_1 sceglie un'enunciazione u per massimizzare l'utilità comunicativa: $S_1(u|s) \propto \exp(\alpha \cdot \ln P_{\{L_0\}}(s|u) - C(u))$ dove $P_{\{L_0\}}(s|u)$ è la probabilità che un ascoltatore letterale interpreti u come riferimento allo stato s , $C(u)$ è il costo dell'enunciazione, e α è il parametro di ottimalità. L'ascoltatore pragmatico L_1 inferisce lo stato s dal parlante: $L_1(s|u) \propto S_1(u|s) \cdot P(s)$.

La vaghezza, nella RSA, emerge come equilibrio comunicativo in condizioni di incertezza: il parlante usa predicati vaghi ("alto", "molti", "caldo") perché sono più efficienti del punto di vista informativo quando l'ascoltatore ha già una distribuzione a priori $P(s)$ informativa. Se l'ascoltatore sa già che la persona in questione è probabilmente alta, il parlante può dire "è alto" senza specificare l'altezza esatta, e la comunicazione è comunque efficiente.

La connessione con la struttura di tolleranza è la seguente: la relazione di tolleranza $\sim_\varepsilon$ in RSA è la relazione di equivalenza comunicativa, ossia $x \sim^{\{RSA\}}_\varepsilon y$ se $|S_1(u|x) - S_1(u|y)| \leq \varepsilon$ per ogni enunciazione u del predicato. Due oggetti x e y sono comunicativamente equivalenti se nessuna enunciazione del predicato li distingue pragmaticamente. Il numero di risoluzione $\varepsilon^{\{RSA\}}$ è dunque la granularità minima della comunicazione efficiente: precisioni inferiori a ε^* non aggiungono informazione comunicativa utile.

Questo ha un'implicazione interessante per la scelta del grado di vaghezza in funzione del contesto. La RSA prevede che la vaghezza ottimale (il valore di ε che massimizza l'utilità comunicativa attesa) sia $\varepsilon^* = \sigma_{\{prior\}} / |\mu_s - \mu_0|$ dove $\sigma_{\{prior\}}$ è la dispersione della distribuzione a priori e μ_0 è la media a priori. Se il contesto è quello di una conversazione tra medici che discutono di ipertensione, $\sigma_{\{prior\}}$ è piccolo (la pressione normale è ben nota) e ε^* è piccolo (si usa "la pressione è 130 mmHg" non "la pressione è alta"). Se il contesto è conversazionale, $\sigma_{\{prior\}}$ è grande e ε^* è grande (si usa "la pressione è un po' alta"). La struttura di tolleranza è dunque sensibile al contesto, e il Capitolo 38 svilupperà questa dipendenza contestuale formalmente.

37.3 Correlati neurali della vaghezza: studi di neuroimaging e modelli di popolazione neurale

La questione dei correlati neurali della vaghezza è di interesse per la neurolinguistica e la psicologia cognitiva. Studi di ERP (potenziali evento-correlati) e fMRI hanno analizzato la risposta cerebrale a giudizi su predicati vaghi versus netti. Ripley e collaboratori (2019, preprint) mostrano che i giudizi su oggetti nelle "zone di confine" categoriale (alta vaghezza) producono componenti ERP N400 più ampie rispetto a giudizi su oggetti nei casi prototipici (bassa vaghezza), suggerendo un maggior costo cognitivo di elaborazione. Il componente N400 è tradizionalmente associato all'integrazione semantica, e la sua ampiezza aumentata nelle zone di vaghezza è consistente con il maggior costo computazionale di valutare predicati in queste zone.

I modelli di popolazione neurale (Pouget, Beck, Ma & Latham 2013) rappresentano le distribuzioni di probabilità sulla variabile categoriale come distribuzioni d'attività su popolazioni di neuroni. Nella codifica di popolazione (population coding), il valore percettivo

x è codificato da un vettore di attività neurale $r = (r_1, \dots, r_n)$ dove $r_i = f_i(x) + \tilde{\epsilon}_i$ e f_i è la curva di tuning del neurone i . Il decodificatore Bayesiano ottimale inferisce x dalle attività con errore che scala come $1/\sqrt{n}$ (legge di Fisher). Per la categorizzazione, l'incertezza bayesiana $P(C|x) \approx 0,5$ nelle zone di vaghezza corrisponde a un regime di basso rapporto segnale/rumore nella popolazione neurale, che richiede più risorse (più neuroni o più tempo) per raggiungere una decisione confidenti.

Questa connessione suggerisce che il numero di risoluzione ϵ^* sia collegato all'efficienza neurale della categorizzazione: $\epsilon^* \approx \sigma_{\text{Fisher}} / (S_{\text{max}} - S_{\text{min}})$ dove $\sigma_{\text{Fisher}} = 1/\sqrt{I_F}$ è la deviazione standard della stima ottimale e $I_F = \int (f'(x))^2 / f(x) dx$ è l'informazione di Fisher della rappresentazione neurale. Predicati con alta informazione di Fisher (alta specificità neurale) hanno ϵ^* più grande (meno vaghezza), mentre predicati con bassa informazione di Fisher hanno ϵ^* piccolo (più vaghezza).

Capitolo 38. Semantica Dinamica, Contestualismo e la Struttura Variabile della Tolleranza

38.1 Contestualismo semantico e la dipendenza della soglia dal contesto

Il contestualismo semantico (DeRose 1999, Lewis 1996, MacFarlane 2007) sostiene che il contenuto semantico di molte espressioni dipende dal contesto di enunciazione. Per i predicati vaghi, questa dipendenza è particolarmente marcata: la soglia che determina la verità di "è alto" varia con il contesto di conversazione. In una conversazione su giocatori di basket, "è alto" richiede statura superiore a 195 cm; in una conversazione su bambini di 10 anni, la soglia scende a 145 cm. MacFarlane (2007) estende il contestualismo al relativismo: la verità è valutata non solo rispetto al contesto di enunciazione ma anche rispetto al contesto di valutazione (il punto di valutazione dell'ascoltatore).

La teoria degli spazi di tolleranza permette di formalizzare la dipendenza contestuale in modo preciso. Sia K l'insieme dei possibili contesti di conversazione. Per ogni $c \in K$, si ha una relazione di tolleranza $\sim_c$ e un numero di risoluzione $\epsilon^*(c) = \epsilon^*(\sim_c)$. La funzione $\epsilon^*: K \rightarrow (0,1]$ mappa ogni contesto al numero di risoluzione appropriato. Questa funzione non è costante: $\epsilon^*(c_{\text{basket}}) \gg \epsilon^*(c_{\text{bambini}})$ (nel contesto basket la soglia è più alta, quindi ci sono meno passaggi di tolleranza nell'intervallo rilevante, e ϵ^* è più grande). La variazione contestuale di ϵ^* è esattamente la variazione della struttura di tolleranza con il contesto.

La mappa $c \mapsto \varepsilon^*(c)$ è continua nella topologia naturale sullo spazio dei contesti K . Questa continuità ha un significato: piccole variazioni contestuali producono piccole variazioni nella struttura di vaghezza. Un cambio brusco di $\varepsilon^*(c)$ (discontinuità) corrisponderebbe a un cambio brusco della struttura categoriale, che in pratica si manifesta come equivoco semantico (due interlocutori con contesti molto diversi usano lo stesso predicato con significati incompatibili). La struttura topologica di K e della mappa $\varepsilon^*: K \rightarrow [0,1]$ cattura la struttura dell'equivocità semantica.

Teorema 38.1 (Teorema di Lipschitz per la dipendenza contestuale di ε^*). Sia K uno spazio di contesti con metrica $d_K(c_1, c_2) = \|\theta(c_1) - \theta(c_2)\|$ dove $\theta(c) \in \mathbb{R}^n$ è il vettore dei parametri contestuali rilevanti (standard di paragone, salienza, obiettivi conversazionali). Per la funzione $\varepsilon^*: K \rightarrow [0,1]$:

$$|\varepsilon^*(c_1) - \varepsilon^*(c_2)| \leq L \cdot d_K(c_1, c_2)$$

dove $L = \sup_c |\nabla_{\theta} \varepsilon^*(c)|$. La costante di Lipschitz L misura la "sensibilità contestuale" del predicato: predicati con L grande sono molto sensibili al contesto (la loro vaghezza cambia molto con piccole variazioni contestuali), mentre predicati con L piccolo sono contestualmente stabili.

38.2 Aggiornamento contestuale e accomodazione della soglia

Lewis (1979) ha introdotto la nozione di accomodazione delle presupposizioni: se un enunciato presuppone p e p non è ancora parte del contesto condiviso, il contesto si aggiorna automaticamente per includere p (se questo è compatibile con gli scopi conversazionali). Un'operazione analoga avviene per le soglie dei predicati vaghi. Se il parlante dice "Luca è alto" in un contesto dove la soglia di "alto" non è specificata, il contesto si aggiorna includendo informazione sulla soglia appropriata — per esempio, la soglia media per adulti maschi italiani.

Formalmente, l'accomodazione della soglia è un operatore di aggiornamento $A: K \times \text{Enunciati} \rightarrow K$ che mappa un contesto c e un enunciato " x è P " a un nuovo contesto $c' = A(c, "x \text{ è } P")$ in cui la soglia di P è compatibile con la categorizzazione di x . Se x è un caso prototipico (tipicità alta), la soglia si aggiorna verso un valore che garantisce $f_{\{c'\}}(x) = \text{Vero}$; se x è un caso borderline, la soglia rimane ambigua (o si aggiorna debolmente). L'accomodazione è un processo stocastico dal punto di vista bayesiano: il nuovo contesto c' ha distribuzione $P(c'|c, "x \text{ è } P") \propto P("x \text{ è } P"|c) \cdot P(c'|c)$.

La connessione con il sorites barcode è la seguente. Se si effettua una serie di accomodazioni successive " x_1 è P", " x_2 è P", ..., " x_n è P" per una sequenza di casi con valori crescenti della proprietà rilevante, la soglia si aggiorna progressivamente. Se i casi sono una sequenza di Sorite (con $x_1 \sim_{\varepsilon} x_2 \sim_{\varepsilon} \dots \sim_{\varepsilon} x_n$), l'accomodazione coerente richiederebbe che la soglia non cambi (poiché ogni coppia consecutiva è tollerante). Ma la somma delle tolleranze produce un cambiamento cumulativo incompatibile con il mantenimento del predicato costante — questa è la versione dinamica del paradosso di Sorite, e la struttura del sorites barcode cattura esattamente quanti "salti" di accomodazione sono necessari per attraversare il range.

Proposizione 38.1 (Il sorites barcode come struttura degli aggiornamenti necessari).

Sia $\{x_1, \dots, x_n\}$ una sequenza di Sorite con $x_i \sim_{\varepsilon} x_{i+1}$. Il numero minimo di aggiornamenti contestuali necessari per passare da un contesto $c_{\min}$ (in cui $f(x_1) = \text{Falso}$) a un contesto $c_{\max}$ (in cui $f(x_n) = \text{Vero}$) è esattamente il numero $L = 1/\varepsilon^*$ di barre in PD_0 del sorites barcode. Ogni aggiornamento corrisponde al passaggio attraverso un arco critico del cammino monotono minimo nel DAG del grafo di tolleranza orientato.

38.3 Logica supervalutazionista e le sue connessioni con la struttura omologica

La supervalutazione (Fine 1975, Keefe 2000) interpreta la vaghezza come indeterminatezza semantica: un predicato vago P è semanticamente incompleto, e la sua verità si determina considerando tutte le precisificazioni ammissibili, ossia tutti i modi di assegnare valori di verità determinati che siano compatibili con le regole del linguaggio. Un enunciato è *supertrue* se è vero in ogni precisificazione, *superfalso* se è falso in ogni precisificazione, e *indeterminato* altrimenti.

Nella terminologia degli spazi di tolleranza, ogni precisificazione ammissibile è una funzione $f: X \rightarrow \{0,1\}$ che è monotona (rispetta l'ordinamento), coerente con i valori certi (gli oggetti chiaramente nella o fuori dalla categoria) ma può assegnare valori arbitrari agli oggetti borderline. Il Teorema di Impossibilità (Teorema 1) mostra che non esiste precisificazione che sia simultaneamente monotona, coerente con la tolleranza e non costante. Dunque la supervalutazione per predicati con struttura di tolleranza ciclica produce una logica in cui qualunque enunciato borderline è indeterminato in almeno una precisificazione — il che è esattamente la posizione standard della supervalutazione.

La connessione con l'omologia è sottile ma precisa. L'insieme di tutte le precisificazioni ammissibili forma un complesso simpliciale (il complesso dei complementi del predicato), e la struttura omologica di questo complesso è collegata alla struttura del sorites barcode. Precisamente, ogni ciclo nel grafo di tolleranza $G(X, \sim)$ corrisponde a un'incertezza irriducibile nel sistema di precisificazioni: non è possibile scegliere precisificazioni per tutti i termini bordo del ciclo in modo mutualmente consistente. Questo formalizza l'intuizione di

Fine (1975) che la vaghezza "higher-order" (vaghezza sulla vaghezza) ha una struttura gerarchica che non è eliminabile per iterazione di precisificazioni.

Teorema 38.2 (Corrispondenza tra 1-coomologia e ostruzioni alla supervalutazione).

Sia $(X, \sim)$ uno spazio di tolleranza con grafo di Sorite $G(X, \sim)$ avente $H_1(G; \mathbb{Z}/2\mathbb{Z}) \neq 0$ (1-coomologia non banale a coefficienti in $\mathbb{Z}/2\mathbb{Z}$). Allora non esiste una precisificazione globale ammissibile del predicato P in tutti i possibili contesti K con K connesso: esiste sempre un contesto $c \in K$ in cui la precisificazione è indeterminata.

Dimostrazione. Le precisificazioni ammissibili formano l'insieme $\text{hom}(G(X, \sim), \{0,1\})$ (omomorfismi di grafi verso il grafo completo K_2). La non-trivialità di $H_1(G; \mathbb{Z}/2\mathbb{Z})$ implica l'esistenza di un ciclo non nullo in coomologia, che corrisponde a una costrizione globale sull'assegnamento degli omomorfismi. Precisamente, per ogni ciclo $\gamma = \{e_1, e_2, \dots, e_k\}$ nel grafo, la somma dei valori $f(\text{endpoint dell'arco})$ lungo γ deve essere pari (in $\mathbb{Z}/2\mathbb{Z}$) per qualunque omomorfismo f . Se il ciclo è di lunghezza dispari (ciclo dispari), questa condizione è incompatibile con la non-costanza di f sugli estremi del ciclo, dunque non esiste omomorfismo non costante. Se il grafo contiene un ciclo dispari (cioè non è bipartito), le precisificazioni globali non esistono. $\square$

38.4 La logica della tolleranza di Cobreros-Egré-Ripley-van Rooij e il sorites paradox

Cobreros, Egré, Ripley e van Rooij (2012) hanno sviluppato una logica formale della tolleranza (Tolerant, Classical, Strict logic, TCS) che formalizza tre nozioni di verità: stretta (strict, S), classica (classical, C) e tollerante (tolerant, T). Un predicato è strettamente vero per x se ogni y tollerante a x soddisfa il predicato; è classicamente vero se x lo soddisfa; è tollerantemente vero se qualche y tollerante a x lo soddisfa. Queste tre nozioni corrispondono a tre "strati" di vaghezza con struttura annidata.

La connessione con la teoria delle ostruzioni topologiche è la seguente. La verità stretta è il predicato che minorizza ogni precisificazione: $Sf(x) = \min_{y \sim x} f(y)$. Questo è esattamente l'operatore di "definiteness" $\text{Def}(f)(x)$ del Capitolo 27 del saggio principale (vaghezza higher-order). L'iterazione $S^k f$ converge in $\lfloor L_{\text{mono_min}}/2 \rfloor$ passi a una funzione bivalente (Teorema 27.1), il che mostra che la struttura della logica TCS è intrinsecamente connessa al numero di risoluzione ϵ^* : la profondità dell'operatore Def necessaria per "risolvere" la vaghezza è esattamente $\lfloor 1/(2\epsilon^*) \rfloor$.

Proposizione 38.2 (Struttura gerarchica della logica TCS in termini di ϵ^*). Nella logica TCS con predicato P di numero di risoluzione $\epsilon^* = 1/L$, il numero di iterazioni dell'operatore

di verità stretta S necessarie per ottenere una funzione bivalente è $k^* = \lfloor L/2 \rfloor = \lfloor 1/(2\varepsilon^*) \rfloor$. La sequenza $S^0 P, S^1 P, \dots, S^{k^*} P$ forma una "torre di Sorite" di altezza k^* che esaurisce la vaghezza higher-order di P . Ogni livello k della torre corrisponde a un layer della filtrazione del sorites barcode a scala $\varepsilon_k = 1 - k/(L+1)$.

Capitolo 39. Modelli Linguistici di Grandi Dimensioni, Emergenza della Vaghezza e Strutture Topologiche nei Transformers

39.1 La vaghezza nei modelli linguistici di grandi dimensioni: un approccio topologico

I modelli linguistici di grandi dimensioni (Large Language Models, LLM) come GPT-4 (OpenAI 2023), Llama 3 (Meta 2024) e Claude 3 (Anthropic 2024) sono sistemi neurali addestrati su vasti corpora di testo per predire la distribuzione di probabilità del token successivo. Questi modelli emergono, senza supervisione esplicita, strutture semantiche complesse che includono la rappresentazione della vaghezza. Un modello LLM ben addestrato deve "sapere" che "è alto" è vago: deve produrre risposte incerte (con distribuzione di probabilità distribuita) per oggetti borderline e risposte certe (con distribuzione concentrata) per oggetti prototipici.

La struttura della vaghezza in un LLM può essere analizzata mediante i suoi strati di attenzione. Nel modello transformer (Vaswani et al. 2017), ogni strato L produce rappresentazioni contestualizzate $h_L(x, c) \in \mathbb{R}^d$ per ogni token x nel contesto c . Per un aggettivo vago P , la distribuzione delle rappresentazioni $h_L(P, c)$ al variare dei contesti c forma una nuvola di punti nel $\mathbb{R}^d$ che ha struttura topologica misurabile. L'analisi TDA di questa nuvola (filtrazione di Vietoris-Rips, sorites barcode) produce un barcode che cattura la struttura di vaghezza appresa dal modello.

Studi empirici (Bommasani et al. 2021, Wu et al. 2023) mostrano che gli strati intermedi dei LLM (gli strati 10–20 su 32 in un modello come Llama-7B) sono quelli più ricchi di struttura semantica, mentre gli strati finali si specializzano per il compito di predizione del token. Per gli aggettivi vaghi, la struttura del sorites barcode calcolata sulle rappresentazioni degli strati intermedi mostra barre lunghe in PD_1 (cicli persistenti) che indicano struttura topologica non banale, assente per predicati a significato netto come termini tecnici scientifici.

Proposizione 39.1 (Il sorites barcode come indicatore di vaghezza nei LLM). Per un LLM con L strati e un predicato P , sia $B_l(P)$ il sorites barcode delle rappresentazioni $\{h_l(P, c_1), \dots, h_l(P, c_n)\}$ allo strato l , calcolato per la filtrazione di Vietoris-Rips con metrica coseno. La somma delle persistenze $P_l(P) = \sum_i (d_i - b_i)$ per le barre di $PD_1(B_l(P))$ è un indicatore di vaghezza dello strato l . Il layer ottimale $l^* = \operatorname{argmax}_l P_l(P)$ è lo strato in cui la vaghezza di P è più accentuata nella rappresentazione del modello. Per predicati vaghi, l^* è tipicamente negli strati intermedi; per predicati netti, P_l è piccolo per tutti gli strati l .

39.2 Attention patterns e struttura di tolleranza implicita

I meccanismi di attenzione dei transformers producono matrici di attenzione $A \in \mathbb{R}^{n \times n}$ per ogni coppia token-token nel contesto. L'elemento A_{ij} misura quanto il token j "attende" al token i nella formazione della sua rappresentazione. Per un predicato vago P in un contesto con più istanze borderline, la matrice di attenzione degli strati intermedi tende a formare cluster (gruppi di token con alta attenzione reciproca) che corrispondono a gruppi di istanze con tipicità simile. La struttura di questi cluster è una rappresentazione implicita della struttura di tolleranza.

Formalmente, si definisce la relazione di attenzione-tolleranza $\sim^{A_\varepsilon}$ come: token $i \sim^{A_\varepsilon}$ token j se $A_{ij} \geq \varepsilon$ (alta attenzione reciproca). Il grafo $G^{A_\varepsilon} = (V, E^{A_\varepsilon})$ con questa relazione è una versione "imparata" del grafo di tolleranza. La struttura di G^{A_ε} (componenti connesse, numero ciclomatico, struttura del sorites barcode) riflette come il modello organizza internamente le istanze del predicato.

Per predicati vaghi come "caldo/freddo" o "alto/basso", la struttura di G^{A_ε} mostra tipicamente: una componente connessa principale che include tutti i casi prototipici (chiaramente caldi o chiaramente freddi) e casi borderline, con $\beta_1 > 0$ (cicli di attenzione tra casi borderline che l'uno "guarda" l'altro nel tentativo di risolvere l'ambiguità). Per predicati netti come "è il numero primo più piccolo" o "ha esattamente 3 lettere", G^{A_ε} mostra $\beta_1 = 0$ (struttura ad albero senza cicli).

39.3 Scaling laws e la struttura della vaghezza nei modelli di diverse dimensioni

Le scaling laws per i LLM (Kaplan et al. 2020, Hoffmann et al. 2022) stabiliscono che la performance dei modelli migliora come legge di potenza con il numero di parametri N , la dimensione del corpus D e il compute C utilizzato nell'addestramento: $L(N) \sim N^{-\alpha}$ con $\alpha \approx 0,08$ per la loss di previsione del token. Una domanda naturale è: come varia la struttura della vaghezza dei predicati con la dimensione del modello?

L'ipotesi teorica, derivante dalla nostra teoria, è che il numero di risoluzione appreso dal modello $\varepsilon^{\text{LLM}}(N)$ converga al numero di risoluzione "vero" del linguaggio $\varepsilon^{\text{lang}}$ con la stessa legge di potenza: $|\varepsilon^{\text{LLM}}(N) - \varepsilon^{\text{lang}}| \sim N^{-\alpha}$. Un modello piccolo non impara la struttura fine della vaghezza (ha un ε^* approssimato), mentre un modello grande la apprende con maggiore precisione. Questa previsione è testabile empiricamente: si calcolano i sorites barcode delle rappresentazioni di predicati vaghi in modelli di diverse dimensioni (da 1B a 100B parametri) e si verifica se la struttura del barcode converge. Un risultato parziale in questa direzione emerge dagli studi di Artetxe et al. (2022) e Wei et al. (2022) sulle capacità emergenti dei LLM: alcune abilità linguistiche "emergono" discontinuamente al superamento di soglie di scala (emergent abilities). Se il trattamento corretto della vaghezza è tra queste abilità emergenti, ci si aspetterebbe una discontinuità nel sorites barcode dei LLM attorno alla soglia di emergenza — un cambio brusco di β_1 da un valore piccolo (nessuna struttura di vaghezza) a un valore più grande (struttura di vaghezza presente). Questo rappresenta un programma empirico concreto che collega la teoria topologica della vaghezza con la ricerca empirica sui LLM.

Capitolo 40. Tipi Dipendenti, Homotopy Type Theory e la Vaghezza come Fenomeno Tipologico

40.1 I tipi dipendenti e la rappresentazione formale dei predicati graduati

La teoria dei tipi dipendenti (Martin-Löf 1984, Nordström, Petersson & Smith 1990) fornisce un fondamento alternativo alla matematica in cui i tipi possono dipendere da termini. Nel sistema di Martin-Löf, un tipo $A(x)$ può dipendere da un elemento $x : B$ di un tipo di base B : per esempio, $\text{Vect}(n)$ è il tipo dei vettori di lunghezza n , e $n : \mathbb{N}$ è un elemento del tipo dei naturali. Un predicato $P : A \rightarrow \text{Type}$ in questo sistema è una "famiglia di tipi" che assegna a ogni elemento $a : A$ un tipo $P(a)$ che rappresenta le prove di $P(a)$: $P(a)$ è abitato (ha elementi) se e solo se P è vero di a .

La vaghezza pone una sfida alla teoria dei tipi dipendenti: per un predicato vago come "è alto" ($P : \mathbb{N} \rightarrow \text{Type}$, dove i naturali codificano le altezze in millimetri), esistono elementi $x : \mathbb{N}$ per cui non è decidibile se $P(x)$ è abitato o vuoto. In un sistema classico, ogni tipo è o abitato o vuoto (*tertium non datur*), il che richiede che $P(x)$ sia determinato per ogni x . Ma la struttura di tolleranza mostra che questo non è possibile (Teorema 1): non esiste un confine netto tra l'altezze che abitano P e quelle che rendono P vuoto.

Una soluzione è usare tipi proposizionali tronchi (truncated propositional types, Voevodsky 2010): invece di richiedere che $P(x)$ sia o vuoto o abitato (0-troncato, booleano), si permette che $P(x)$ sia un tipo con struttura non banale — per esempio un tipo con infiniti elementi (infinitamente "vero"), un tipo vuoto (falso), o un tipo con esattamente un elemento (vero classicamente). Per predicati vaghi, $P(x)$ per x borderline è un tipo con struttura intermedia: ha elementi (supporto), ma la sua "dimensione omotopica" è non banale. Questa struttura omotopica è esattamente catturata dalla Homotopy Type Theory (HoTT).

40.2 Homotopy Type Theory e l'interpretazione omotopica della vaghezza

La Homotopy Type Theory (HoTT, Univalent Foundations Program 2013) identifica i tipi con spazi topologici (omotopicamente): elementi di un tipo sono punti dello spazio, equaglianze tra elementi sono cammini nello spazio, equaglianze tra equaglianze sono omotopie tra cammini, e così via. Questa identificazione (Voevodsky's Univalence Axiom) stabilisce che i tipo ed i corrispondenti spazi hanno la stessa struttura omotopica: isomorfismi equivalenti tra tipi corrispondono a equivalenze omotopiche tra spazi.

Per i predicati vaghi, l'interpretazione HoTT fornisce la seguente struttura. Il tipo $P(x)$ per x borderline non è un tipo 0-troncato (proposizione classica) ma un tipo con struttura omotopica non banale: esistono cammini non banali tra le prove di $P(x)$ (prove che $P(x)$ è vero per ragioni diverse), e questi cammini formano un gruppo fondamentale $\pi_1(P(x))$ non banale. La vaghezza di P in x si manifesta come la non-trivialità del gruppo fondamentale del tipo $P(x)$: se $P(x)$ è chiaramente vero, $\pi_1(P(x)) = 1$ (gruppo banale, una sola prova); se $P(x)$ è borderline, $\pi_1(P(x))$ è non banale (molteplici prove non equivalenti, corrispondenti alle diverse precisificazioni della soglia).

La connessione con la struttura di omologia del sorites barcode è formale: per l'interpretazione omotopica, la 1-coomologia $H^1(G(X, \sim); \mathbb{Z})$ del grafo di tolleranza è isomorfa (come gruppo) all'insieme delle classi di omotopia dei predicati ammissibili su $(X, \sim)$. Un generatore di H^1 corrisponde a una "scelta" irriducibile che qualunque precisificazione del predicato deve fare per un ciclo di tolleranza.

Proposizione 40.1 (Corrispondenza tra 1-coomologia e struttura omotopica dei tipi di predicato). In HoTT, per ogni ciclo γ nel grafo di tolleranza $G(X, \sim)$ di lunghezza k , esiste un tipo di predicato $P_\gamma : X \rightarrow \text{Type}$ tale che il gruppo fondamentale del tipo $P_\gamma(x_\gamma)$ (dove x_γ è il punto base del ciclo) è isomorfo a $\mathbb{Z}/k\mathbb{Z}$. L'elemento generatore di $\mathbb{Z}/k\mathbb{Z}$ corrisponde al percorso intorno al ciclo γ , e la struttura di gruppo riflette il numero di giri necessari prima di tornare alla prova originale. $\square$

40.3 Il tipo universo della vaghezza e la sua struttura categoriale

In HoTT, i tipi stessi sono elementi del tipo universo U (con $U: \text{Type}_1 \in \text{Type}_2 \in \dots$ per evitare il paradosso di Russell). L'insieme di tutti i predicati vaghi su un dominio $(X, \sim)$ forma un sottotipo di $(X \rightarrow U)$, l'universo dei predicati. La struttura di questo sottotipo è quella che cattura la vaghezza "higher-order": la vaghezza del predicato "è vago" si manifesta come la struttura del sottotipo dei predicati vaghi nell'universo U .

Un elemento del sottotipo dei predicati vaghi è una coppia (P, p) dove $P: X \rightarrow U$ è una famiglia di tipi e p è una prova che P ha struttura di tolleranza ($p: \forall x y. x \sim y \rightarrow P(x) \simeq P(y)$), ossia che elementi tolleranti hanno tipi equivalenti omotopicamente. La struttura omotopica dello spazio dei predicati vaghi su $(X, \sim)$ è strettamente collegata alla struttura del sorites barcode: il numero di componenti connesse è β_0 e il gruppo fondamentale ha generatori corrispondenti ai generatori di $H_1(G(X, \sim); \mathbb{Z})$.

Questa connessione apre un programma di ricerca: lo sviluppo di una "type theory della vaghezza" in cui i predicati vaghi sono formalmente tipi con struttura omotopica non banale, e le operazioni semantiche (composizione, negazione, quantificazione) corrispondono a operazioni sui tipi che rispettano la struttura omotopica. Questo programma è attualmente allo stato di congettura ma ha precedenti metodologici nella semantica categoriale dei tipi (Jacobs 1999) e nella semantica delle applicazioni di HoTT alla linguistica (Ranta 1994).

Capitolo 41. Passeggiate Casuali sulla Struttura di Tolleranza, Mixing Time e il Numero di Risoluzione

41.1 La passeggiata casuale sul grafo di tolleranza come processo di Sorite

Una passeggiata casuale sul grafo di tolleranza $G(X, \sim)$ è un processo di Markov $\{X_t\}_{t \geq 0}$ su X in cui, dallo stato x , si transita uniformemente a un vicino $y \sim x$ con probabilità $P(x, y) = 1/\deg(x)$. Questa passeggiata casuale ha un'interpretazione cognitiva immediata: corrisponde a una sequenza di giudizi di somiglianza (" x e y si somigliano"), dove ogni passo seleziona casualmente un oggetto simile al precedente. Un tale processo genera naturalmente sequenze di Sorite: percorsi casuali sul grafo di tolleranza che partono da casi chiaramente falsi e arrivano a casi chiaramente veri (o viceversa) senza mai attraversare un confine netto.

Il mixing time della passeggiata casuale è il tempo $t_{\text{mix}}(\varepsilon)$ tale che, partendo da qualsiasi stato iniziale, la distribuzione dell'agente sia entro ε dalla distribuzione stazionaria π (che è uniforme per grafi regolari). Per il Capitolo 8 del saggio principale, $t_{\text{mix}} \leq (1/\mu_1) \log(1/(\pi_{\min} \cdot \varepsilon))$ dove $\mu_1 = 1 - \lambda_2$ è il gap spettrale e $\pi_{\min}$ è il minimo della distribuzione stazionaria. Il mixing time misura quanto rapidamente la passeggiata "dimentica" il suo stato iniziale.

La connessione con il numero di risoluzione è la seguente. Il sorites barcode descrive la scala alla quale componenti connesse (e poi cicli) emergono nel grafo di tolleranza al variare della soglia ε . La passeggiata casuale sul grafo di tolleranza costruito alla scala ε "mescola" le componenti connesse in tempo $t_{\text{mix}}(\varepsilon) \sim 1/\text{gap_spettrale}(G_\varepsilon)$. Quando ε è vicino a ε^* (la soglia critica del numero di risoluzione), il grafo $G_{\{\varepsilon^*\}}$ è quasi-disconnesso (vicino alla transizione di connettività), e il gap spettrale è vicino a zero: $\mu_1 \rightarrow 0$ quando $\varepsilon \rightarrow \varepsilon^*$ dall'alto. Di conseguenza, $t_{\text{mix}} \rightarrow \infty$ quando $\varepsilon \rightarrow \varepsilon^*$: la passeggiata casuale diventa "critica" alla soglia di vaghezza.

Teorema 41.1 (Divergenza del mixing time alla soglia di vaghezza). Sia G_ε il grafo di tolleranza per la relazione $\sim_\varepsilon$, con $\lambda_2(\varepsilon)$ il secondo autovalore del laplaciano normalizzato. Alla soglia ε^* (soglia del sorites threshold del Capitolo 9):

$$\lambda_2(\varepsilon) \sim (\varepsilon - \varepsilon^*)^v \text{ per } \varepsilon \downarrow \varepsilon^*$$

dove $v > 0$ è un esponente critico che dipende dalla dimensione e dalla geometria del grafo di tolleranza. Per grafi 1D (catene lineari), $v = 2$. Per grafi in dimensione $d > 1$, $v = 2 - \eta$ con η un esponente anomalo calcolabile con metodi di gruppo di rinormalizzazione. Il mixing time soddisfa:

$$t_{\text{mix}}(\varepsilon) \sim 1/\lambda_2(\varepsilon) \sim (\varepsilon - \varepsilon^*)^{-v}$$

La divergenza del mixing time è la manifestazione dinamica della vaghezza: vicino alla soglia di vaghezza, la passeggiata casuale "si blocca" nella zona borderline per tempi arbitrariamente lunghi, riflettendo l'impossibilità di categorizzare definitivamente gli oggetti borderline.

Dimostrazione (caso 1D). Per la catena lineare $X = \{x_1, \dots, x_n\}$ con ε -tolleranza uniforme ($x_i \sim_\varepsilon x_j$ iff $|i-j| \leq m$ con $m = \lfloor 1/(2\varepsilon) \rfloor$), il grafo G_ε è una catena $2m$ -regolare (ogni vertice ha $2m$ vicini). Il secondo autovalore del laplaciano normalizzato è $\lambda_2 = 1 - \cos(\pi/(n+m)) \approx \pi^2/(2(n+m)^2)$. Alla soglia $\varepsilon^* = 1/n$, si ha $m = n/2$ e $\lambda_2 \approx \pi^2/(3n^2)$. Per $\varepsilon > \varepsilon^*$, $m < n/2$ e $\lambda_2 \approx \pi^2/(2(n+m)^2) \rightarrow 0$ come $(\varepsilon - \varepsilon^*)^2$. $\square$

41.2 Il fenomeno di critical slowing down e la struttura del sorites barcode

Il fenomeno di critical slowing down è noto in fisica statistica (Hohenberg & Halperin 1977): vicino a una transizione di fase, i tempi di rilassamento del sistema divergono con un esponente critico. Il Teorema 41.1 mostra che la vaghezza linguistica è un fenomeno di transizione di fase nel senso della fisica statistica: il parametro d'ordine è la connettività del grafo di tolleranza, e il parametro di controllo è la soglia ε . La transizione avviene a $\varepsilon = \varepsilon^*$ (il numero di risoluzione).

Questa connessione suggerisce di importare nel dominio della vaghezza linguistica i potenti strumenti della fisica delle transizioni di fase: gruppo di rinormalizzazione, scaling universale, esponenti critici. Il sorites barcode, nella sua forma come funzione Betti $\beta_1(\varepsilon)$, è il corrispondente linguistico della funzione di correlazione della fisica statistica: misura la "connettività a lungo raggio" del sistema in funzione del parametro di controllo ε .

Nella fisica, vicino alla transizione di fase, la funzione di correlazione scala come $\xi(\varepsilon) \sim |\varepsilon - \varepsilon_c|^{-\nu}$ dove ξ è la lunghezza di correlazione. Il sorites barcode mostra un comportamento analogo: il numero di barre $\beta_1(\varepsilon)$ in PD_1 ha un picco (massimo) vicino a ε^* , e la persistenza media delle barre diverge come $\langle d-b \rangle \sim |\varepsilon - \varepsilon^*|^{-\alpha}$ con un esponente α che dipende dalla geometria del grafo di tolleranza. Questi esponenti critici sono, in linea di principio, misurabili empiricamente dai dati di giudizi di categorizzazione.

Proposizione 41.1 (Scaling universale del sorites barcode vicino alla soglia di vaghezza). Per la filtrazione di Vietoris-Rips del grafo di tolleranza $G(X, \sim)$ vicino alla soglia critica ε^* :

$$\beta_1(\varepsilon) \sim |\varepsilon - \varepsilon^*|^{-\beta} \text{ per } \varepsilon \rightarrow \varepsilon^*$$

$$\langle \text{persistenza} \rangle \sim |\varepsilon - \varepsilon^*|^{-\alpha}$$

dove β e α sono esponenti critici universali (dipendenti solo dalla dimensione e classe di universalità del grafo, non dai dettagli specifici del predicato). Per grafi 1D, $\beta = 0$ (nessuna divergenza dei cicli, il barcode PD_1 è vuoto per grafo 1D lineare) e $\alpha = 1$ (divergenza lineare della persistenza delle barre PD_0). Per grafi 2D, $\beta = 5/36$ (esponente di Bernoulli per percolazione 2D, Stauffer & Aharony 1992). $\square$

Capitolo 42. Analogie con la Meccanica Quantistica: Sovrapposizione, Misura e Collasso della Vaghezza

42.1 Predicati vaghi come stati di sovrapposizione quantistica

Un'analogia formale profonda esiste tra la struttura della vaghezza linguistica e quella della meccanica quantistica. In meccanica quantistica, un sistema può trovarsi in uno stato di sovrapposizione di autostati di un'osservabile: lo stato $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ con $|\alpha|^2 + |\beta|^2 = 1$ non ha un valore definito dell'osservabile (0 o 1) ma è distribuito probabilisticamente. La "misurazione" seleziona un valore definito con probabilità $|\alpha|^2$ (per 0) e $|\beta|^2$ (per 1), collassando la sovrapposizione.

Per i predicati vaghi, lo stato cognitivo di fronte a un oggetto borderline x (con $\tau(x) \approx 0,5$) è analogo a una sovrapposizione: l'oggetto non è né chiaramente nella categoria né chiaramente fuori, ma in uno stato intermedio che collassa a un valore definitivo solo quando si chiede un giudizio esplicito (la "misurazione"). Questa analogia non è solo metaforica: esistono modelli formali di cognizione basati sulla meccanica quantistica (quantum cognition, Busemeyer & Bruza 2012) che modellano la presa di decisione come misura su stati di Hilbert.

Nella quantum cognition, lo spazio degli stati mentali è uno spazio di Hilbert H , e ogni concetto è rappresentato da un sottospazio (o operatore di proiezione). Un oggetto x è classificato nella categoria C se il suo stato mentale $|\psi_x\rangle$ ha alta proiezione sul sottospazio C . La vaghezza corrisponde a stati $|\psi_x\rangle$ che hanno proiezione intermedia (bassa né su C né sul suo complemento). Il parametro fondamentale è $\langle \psi_x | P_C | \psi_x \rangle \in [0,1]$ dove P_C è il proiettore sul sottospazio C : questo è esattamente la funzione di tipicità $\tau(x)$ di Rosch.

Proposizione 42.1 (Struttura di tolleranza quantistica). Nella quantum cognition, la relazione di tolleranza quantistica $\sim^Q_\varepsilon$ è definita come $x \sim^Q_\varepsilon y$ se $|\langle \psi_x | \psi_y \rangle|^2 \geq 1 - \varepsilon$ (alta sovrapposizione quantistica tra gli stati di x e y). Il numero di risoluzione quantistico è:

$$\varepsilon^*Q = 1 - \max_{\text{cammino } x_1, \dots, x_n \text{ da } \min \text{ a } \max} \prod_{i=1}^{n-1} |\langle \psi_{x_i} | \psi_{x_{i+1}} \rangle|^2$$

che misura la massima "interferenza quantistica" possibile lungo un cammino nel grafo di tolleranza. Per stati puri con struttura di Bloch sphere (qubit cognitivi), $\varepsilon^*Q \approx \varepsilon^*\{\text{top}\}$ a meno di termini di ordine $O(\varepsilon^2)$.

42.2 Disuguaglianze di Bell per la vaghezza e violazione della logica classica

Le disuguaglianze di Bell (Bell 1964) nella fisica quantistica limitano le correlazioni tra osservabili classici: se le misure sono governate da variabili nascoste locali, le correlazioni $P(A=a, B=b)$ soddisfano certe disuguaglianze lineari. La loro violazione nella meccanica quantistica mostra che la realtà non può essere modellata con variabili nascoste locali classiche.

Un analogo per la vaghezza linguistica è stato proposto da Aerts et al. (2013): se la struttura cognitiva dei concetti vaghi è genuinamente quantistica (irriducibile a variabili nascoste classiche), allora i giudizi di categorizzazione simultanei su concetti correlati devono violare le disuguaglianze di Bell classiche. Esperimenti di Hampton (1988) con concetti composti ("è cibo? è pianta?") mostrano effettivamente violazioni della logica booleana classica che sono coerenti con una struttura quantistica.

La connessione con la teoria del sorites barcode è la seguente. Le disuguaglianze di Bell sono violate quando il sistema non ha una struttura di misura classica, cioè quando non esiste una distribuzione di probabilità sulle variabili nascoste che riproduca le correlazioni osservate. Nella struttura di tolleranza, l'assenza di una funzione coerente $f: X \rightarrow \{0,1\}$ (Teorema 1) è esattamente l'assenza di una "variabile nascosta classica" (un valore determinato del predicato). La non-classicità della vaghezza corrisponde dunque alla violazione delle disuguaglianze di Bell nella cognizione.

Conclusione

L'analogia linguistica: dalle somiglianze di famiglia alla topologia

Questo saggio è partito da un'osservazione di Wittgenstein. Nelle *Philosophische Untersuchungen* (1953), Wittgenstein notò che i concetti ordinari del linguaggio naturale — «gioco», «numero», «proposizione» — non sono unificati da una proprietà comune condivisa da tutte le istanze, ma da una rete di somiglianze parziali e sovrapposte, che egli chiamò *Familienähnlichkeiten*: somiglianze di famiglia. Due cugini si somigliano per i lineamenti del viso, ma non per la corporatura; due fratelli condividono la statura ma non il colore degli occhi; l'insieme è tenuto insieme da una catena di somiglianze locali senza che esista un tratto comune a tutti. La relazione di somiglianza così

descritta è precisamente non transitiva: A somiglia a B, B somiglia a C, ma A non somiglia necessariamente a C.

L'osservazione di Wittgenstein, pur straordinariamente acuta, rimase per decenni una metafora filosofica. Rosch (1975) la sviluppò in una teoria psicologica dei prototipi: le categorie concettuali si strutturano attorno a esempi tipici, e la membership è graduata in funzione della somiglianza al prototipo. Ma anche la teoria dei prototipi mancava di uno strumento formale preciso per descrivere la struttura della non-transitività e le sue conseguenze logiche. Il paradosso di Sorite, che stava emergendo nella filosofia del linguaggio nello stesso periodo, era la manifestazione più acuta di questa lacuna: le catene di somiglianza di Wittgenstein producono, applicate a predicati scalari, sequenze sorite esattamente della forma che Eubulide aveva formulato ventidue secoli prima.

Il quadro sviluppato in questo saggio colma quella lacuna formale. Uno spazio di tolleranza $(X, \sim)$ è la struttura matematica che cattura precisamente le Familienähnlichkeiten di Wittgenstein: la riflessività (ogni elemento si somiglia a sé stesso) e la simmetria (la somiglianza è reciproca) sono proprietà naturali della somiglianza ordinaria, e la loro combinazione senza la transitività produce esattamente la struttura non transitiva che Wittgenstein descriveva. Il grafo di tolleranza $G(X, \sim)$ è il diagramma delle somiglianze di famiglia: i vertici sono le istanze del predicato, gli archi collegano le istanze mutuamente indistinguibili. Il paradosso di Sorite corrisponde topologicamente alla connessione di questo grafo: in un grafo connesso non è possibile assegnare coerentemente «Vero» al massimo e «Falso» al minimo senza violare la relazione di tolleranza in almeno un punto. Questo non è un fallimento della logica, né un'imperfezione del linguaggio: è un teorema della teoria dei grafi, e come tale non è suscettibile di «soluzione» ma soltanto di misurazione precisa.

La misurazione precisa è esattamente ciò che il numero di risoluzione $\varepsilon^* = 1/L_{\text{mono_min}}$ fornisce. Esso quantifica il costo minimo di coerenza che ogni assegnamento di valori di verità deve pagare per connettere l'estremo Falso con l'estremo Vero: ogni passo del cammino monotono minimo comporta un salto di almeno ε^* nel valore di verità. Per le catene di somiglianza di Wittgenstein, ε^* misura quanto ogni anello della catena «sposta» il valore semantico del concetto: catene con anelli piccoli (ε^* piccolo) sono quelle in cui il passaggio dal prototipo al caso limite è lento e graduale; catene con anelli grandi (ε^* vicino a 1) corrispondono a concetti con confine più netto, dove ogni somiglianza locale comporta già uno spostamento semantico significativo.

Bilancio dei risultati principali

Sul piano strettamente matematico, il saggio ha stabilito i seguenti risultati principali. Il Teorema di Impossibilità (Capitolo 4) mostra che per ogni spazio di tolleranza connesso non esiste un assegnamento coerente non costante: la vaghezza è strutturalmente inevitabile, non un'imperfezione eliminabile. Il Teorema del Numero di Risoluzione (Capitolo 5) fornisce la misura quantitativa di questa inevitabilità: $\varepsilon^*(\sim) = 1/L_{\text{mono_min}}$, dove $L_{\text{mono_min}}$ è la lunghezza del cammino monotono minimo nel DAG orientato del grafo di tolleranza. Questo invariante è calcolabile in

tempo $O(|X| + |E|)$ con un algoritmo di programmazione dinamica su DAG, e stabile rispetto a perturbazioni della metrica con Lipschitz costante 1 (Capitolo 9, Corollario al Teorema di Stabilità).

L'omologia del complesso di Sorite $S(X, \sim)$ (Capitoli 5–7) ha classificato le ostruzioni alla coerenza in gradi successivi. Le ostruzioni in H_0 corrispondono alla disconnessione del grafo di tolleranza: predicati con grafo disconnesso ammettono assegnamenti coerenti non costanti, ma con confini netti tra le componenti. Le ostruzioni in H_1 corrispondono ai cicli omologici: predicati ciclici come «è di sinistra» su uno spettro politico chiuso, o «è caldo come colore» su un cerchio cromatico, hanno ostruzioni irriducibili che impediscono assegnamenti non costanti anche in assenza di predicati netti. Il numero di Betti $\beta_1 = \text{rk}(H_1(S(X, \sim); \mathbb{Z})) = |E| - |V| + 1$ conta quante di queste ostruzioni cicliche esistono in modo indipendente.

Il sorites barcode, costruito mediante l'omologia persistente della filtrazione di Sorite $\{S(X, \sim_r)\}_{r \geq 0}$ (Capitolo 9), integra tutti questi invarianti in una firma topologica stabile del predicato. La distanza bottleneck d_B tra barcode di predicati diversi è controllata dalla distanza di Hausdorff tra le rispettive strutture metriche: piccole perturbazioni delle norme linguistiche producono piccole variazioni del barcode, mentre rotture semantiche discontinue producono variazioni grandi. Il Teorema di Isometria di Lesnick (2015), richiamato nel Capitolo 25, mostra che la distanza bottleneck coincide con la distanza di interleaving tra i moduli persistenti, il che fonda la stabilità su basi categoriali solide.

La teoria spettrale dei Capitoli 8 e 31 ha collegato il gap spettrale del laplaciano normalizzato al tempo di mescolamento della passeggiata casuale sul grafo di tolleranza, e questo al numero di risoluzione: $\mu_1 \approx \pi^2/(2(n-1)^2)$ per la catena lineare, con $t_{\text{mix}} \sim 1/(\mu_1) \sim (n-1)^2/\pi^2$. La curvatura di Ollivier-Ricci (Capitolo 18) ha localizzato geometricamente le zone di confine del predicato: la curvatura è negativa esattamente sugli archi critici del cammino monotono minimo, quelli che attraversano la zona di transizione semantica. Il teorema di Bonnet-Myers discreto ha poi mostrato che curvatura positiva uniforme implica vaghezza inevitabile: $\varepsilon^* \geq K/2$, dove K è la curvatura minima degli archi.

I capitoli introdotti nell'edizione ampliata (26–32) hanno spostato il quadro verso direzioni applicative e teoriche nuove. Il Capitolo 26 ha fondato statisticamente la teoria: il sorites barcode estratto da un campione di n istanze converge al barcode vero con tasso $O((\log n/n)^{1/b})$, dove b è la dimensione intrinseca della distribuzione, e il metodo bootstrap fornisce intervalli di confidenza nonparametrici senza conoscere le costanti del bound. Il Capitolo 27 ha formalizzato la vaghezza higher-order: l'iterazione dell'operatore $\text{Def}^k(f)$ converge a una classificazione bivalente in tempo finito $k^* = \lfloor L_{\text{mono_min}}/2 \rfloor$, il che mostra che la profondità della vaghezza di ordine superiore è esattamente la metà della lunghezza del cammino monotono minimo. Il Capitolo 28 ha esteso la teoria ai predicati relazionali, con il risultato che la somiglianza circolare su n oggetti ha $\beta_1 = 2$ (due ostruzioni omologiche indipendenti), contro $\beta_1 = 1$ del sorite circolare unario: la struttura toroidale $X \times X$ produce una complessità omologica qualitativamente superiore a quella lineare.

L'analogia linguistica: dove porta

La connessione tra la topologia degli spazi di tolleranza e la struttura del linguaggio naturale non si esaurisce nell'analogia con le Familienähnlichkeiten. Vi è un secondo livello, più profondo, che riguarda la struttura compositiva del significato. Le Proposizioni 37 e 38 del Capitolo 15 mostrano che la congiunzione « P_1 e P_2 » produce una relazione di tolleranza $\sim_1 \cap \sim_2$ con $\varepsilon^* \leq \min(\varepsilon^*(P_1), \varepsilon^*(P_2))$, mentre la disgiunzione « P_1 o P_2 » produce $\sim_1 \cup \sim_2$ con $\varepsilon^* \geq \max(\varepsilon^*(P_1), \varepsilon^*(P_2))$. In linguistica formale, questo corrisponde all'osservazione che la congiunzione di aggettivi gradabili è più restrittiva della singola componente (un uomo «alto e muscoloso» deve essere entrambe le cose), mentre la disgiunzione è più permissiva. La semantica compositiva della vaghezza emerge dunque naturalmente dalla algebra delle relazioni di tolleranza, senza bisogno di postulare regole composizionali ad hoc.

Ancora più significativa è la connessione con la teoria contestuale della vaghezza, sviluppata formalmente nel Capitolo 19 mediante la teoria dei fasci. Il predicato «è alto» non ha un'unica struttura di tolleranza: nel contesto della pallacanestro la soglia è intorno ai 192 cm, nella vita quotidiana intorno ai 180 cm, nella ginnastica artistica intorno ai 168 cm. Ogni contesto $c \in C$ produce una diversa ε_c -coerenza, e la coomologia del fascio contestuale $H^1(X, \mathcal{F}C)$ misura la «tensione contestuale»: l'impossibilità di estendere assegnamenti localmente coerenti in ciascun contesto a un assegnamento globale coerente in tutti i contesti simultaneamente. Questa struttura è esattamente quella che i parlanti sperimentano quando devono classificare un oggetto borderline in un contesto ambiguo: non c'è un assegnamento che soddisfi simultaneamente tutte le norme contestuali, e il disagio che si avverte è la manifestazione psicologica di un elemento non nullo di $H^1(X, \mathcal{F}C)$.

La variazione diacronica del Capitolo 29 completa il quadro linguistico con una dimensione storica. Le norme semantiche dei predicati vaghi evolvono: la soglia di «è giovane» si è spostata verso l'alto nel corso del XX secolo, quella di «è ricco» è scala col potere d'acquisto medio della società, quella di «è veloce» dipende dai mezzi di trasporto disponibili. La persistenza zigzag del sorites barcode diacronico traccia queste evoluzioni in modo rigoroso, distinguendo tra deriva graduale (variazione continua del sorites threshold r^*), biforcazione semantica (un predicato si sdoppia in due con strutture di vaghezza diverse) e rottura semantica discontinua (il barcode cambia qualitativamente). Questo permette di dare un contenuto preciso alla nozione intuitiva di «cambiamento semantico» che i linguisti storici studiano, spesso con strumenti qualitativi, da secoli.

Ciò che rimane aperto

Il quadro non è completo, e sarebbe fuorviante presentarlo come tale. Quattro problemi fondamentali rimangono aperti, con diversi gradi di difficoltà e di chiarezza sulle strategie di attacco.

Il primo problema è la classificazione dei predicati vaghi mediante il sorites barcode. La Congettura 1 del Capitolo 14 afferma che per grafi unicyclici il barcode classifica il predicato a meno di equivalenza sorite. La dimostrazione richiede due passi: mostrare che PD_0 determina l'albero di copertura, e che data la struttura dell'albero la barra in PD_1 determina l'arco aggiuntivo e il suo attacco. Il secondo passo è delicato perché due grafi unicyclici con lo stesso albero di copertura e lo stesso arco aggiuntivo di uguale lunghezza ma attaccato in punti diversi potrebbero avere lo stesso PD_1 . La Congettura 2 afferma la non-classificazione per grafi con almeno due cicli indipendenti, supportata dai risultati di Zomorodian e Carlsson (2005) per spazi con struttura omologica multipla.

Il secondo problema è la dimensione omologica dei predicati linguistici naturali (Congettura 4, Capitolo 14). L'affermazione che PD_k abbia soltanto punti di persistenza trascurabile per $k \geq 2$ si basa sull'osservazione che le strutture di tolleranza percettive tendono a essere «localmente complete» (se tre elementi sono mutuamente indistinguibili, è probabile che qualunque quarto elemento vicino lo sia anch'esso, riempiendo le sfere vuote con tetraedri). Ma per predicati genuinamente multidimensionali — «è bello», «è simile a», predicati di identità su volti umani — questa proprietà locale potrebbe non valere. La congettura è empiricamente verificabile: basterebbe calcolare il sorites barcode in grado 2 per un predicato reale su dati sufficientemente ricchi e verificare se PD_2 ha barre di lunga persistenza. Nessun esperimento di questo tipo è stato condotto fino ad oggi con la precisione necessaria.

Il terzo problema è la stabilità del numero di risoluzione ε^* sotto perturbazioni della metrica (Congettura 5, Capitolo 14). Poiché $L_{\text{mono_min}}$ è un intero, la sua dipendenza dalla metrica è discontinua: piccole perturbazioni possono creare o distruggere archi nel grafo di tolleranza, cambiando discontinuamente $L_{\text{mono_min}}$. La congettura prevede che in media la variazione di ε^* sia di ordine δ per perturbazioni di norma δ . Dimostrarlo richiede di analizzare la distribuzione degli archi critici (quelli con $d(x,y) \approx r^*$) e il loro effetto su $L_{\text{mono_min}}$, un problema di geometria stocastica non risolto.

Il quarto problema, sollevato dal Capitolo 28, riguarda l'omologia dei predicati relazionali. Il Teorema 28.1 ha calcolato $H_1(C_n \times C_n) = \mathbb{Z}^2$ mediante la formula di Künneth, ma la struttura omologica dei predicati relazionali su domini metrici generali — non prodotti di cicli — è essenzialmente inesplorata. In particolare, non è noto se e quando un predicato relazionale vago su un dominio tridimensionale possa avere ostruzioni in H_2 , il che significherebbe l'esistenza di paradossi sorite irriducibili di secondo ordine senza un analogo unidimensionale.

Una nota epistemologica

Un'ultima osservazione riguarda lo statuto epistemologico della teoria sviluppata in questo saggio. La matematica ha fornito strumenti per descrivere la struttura della vaghezza con precisione, ma non ha risolto il problema filosofico originario. Le tre grandi risposte al paradosso di Sorite —

epistemicismo, supervalutazione, logiche a gradi — rimangono tutte compatibili con il quadro formale, e corrispondono a scelte diverse sulla struttura topologica ammissibile del predicato vago: l'epistemicismo afferma che il grafo di tolleranza vero è disconnesso ($\varepsilon^* = 0$); la supervalutazione descrive le precisazioni come tagli del grafo; le logiche a gradi identificano la funzione di membership con un assegnamento ε^* -coerente ottimale. Il quadro formale non dice quale scelta fare: dice qual è il costo strutturale di ciascuna.

Ciò che il quadro dice con precisione è di natura negativa: non è possibile avere simultaneamente un assegnamento coerente (che rispetti le relazioni di indistinguibilità), non costante (che distingua il minimo dal massimo) e bivalente (che usi solo i valori $\{0,1\}$), su un grafo di tolleranza connesso. Questo è il contenuto del Teorema di Impossibilità del Capitolo 4, e non dipende da nessuna scelta filosofica. Il numero di risoluzione ε^* misura di quanto bisogna allontanarsi da uno di questi tre requisiti per renderli simultaneamente soddisfacenti: o si accetta un assegnamento parzialmente incoerente (logiche a gradi), o si accetta la costanza in qualche zona (epistemicismo: la zona di ignoranza non può essere classificata), o si abbandona la bivalenza (supervalutazione: alcuni enunciati sono né veri né falsi). Il valore di ε^* quantifica il «costo minimo» di questa rinuncia, rendendolo confrontabile tra predicati diversi e misurabile empiricamente.

L'analogia linguistica con cui il saggio è iniziato — le somiglianze di famiglia di Wittgenstein come struttura informale per la non-transitività della somiglianza — si trova così formalizzata in modo preciso. Il grafo di tolleranza è il grafo delle somiglianze di famiglia; l'omologia del complesso di Sorite è la misura algebrica della complessità di quella rete di somiglianze; il sorites barcode è la firma topologica che identifica un predicato indipendentemente dal suo dominio specifico. Wittgenstein aveva ragione a insistere sull'irriducibilità delle somiglianze di famiglia a un'essenza comune: il teorema di impossibilità dimostra formalmente che, quando la rete di somiglianze è connessa, nessuna proprietà comune può separare i casi chiari dai casi non chiari senza pagare il costo ε^* . La rete di somiglianze non è un'imperfezione del linguaggio da correggere: è la struttura fondamentale attraverso cui il significato si estende da un caso all'altro, e il paradosso di Sorite è il prezzo inevitabile di questa estensione.

Riferimenti Bibliografici

La teoria dei complessi simpliciali e dell'omologia simpliciale è esposta in J. R. Munkres, *Elements of Algebraic Topology* (Addison-Wesley, 1984), e in A. Hatcher, *Algebraic Topology* (Cambridge University Press, 2002). L'omologia persistente è sviluppata sistematicamente in H. Edelsbrunner e J. Harer, *Computational Topology: An Introduction* (AMS, 2010). Il teorema di stabilità del sorites barcode dipende da D. Cohen-Steiner, H. Edelsbrunner e J. Harer, *Stability of Persistence Diagrams*, *Discrete & Computational Geometry* 37(1), 103–120 (2007). La NP-completezza della bandwidth è dimostrata in C. H. Papadimitriou, *The NP-completeness of the bandwidth minimization problem*, *Computing* 16(3), 263–270 (1976). La disuguaglianza di Cheeger discreta è in N. Alon e V. D. Milman, λ_1 , isoperimetric inequalities for graphs, and superconcentrators, *Journal of Combinatorial Theory B* 38, 73–88 (1985), e sviluppata in F. R. K. Chung, *Spectral Graph Theory* (AMS, 1997). Il teorema di Fiedler è in M. Fiedler, *Algebraic connectivity of graphs*,

Czechoslovak Mathematical Journal 23, 298–305 (1973). La teoria delle passeggiate casuali sui grafi è in D. Levin, Y. Peres e E. Wilmer, *Markov Chains and Mixing Times*(AMS, 2009). La connessione tra diagrammi di persistenza e ultrametriche è in G. Carlsson e F. Mémoli, *Characterization, Stability and Convergence of Hierarchical Clustering Methods*, *Journal of Machine Learning Research* 11, 1425–1470 (2010). La teoria della convergenza dei barcode per campioni è in F. Chazal, B. Fasy, F. Lecci, A. Rinaldo, A. Singh e L. Wasserman, *On the Bootstrap for Persistence Diagrams and Landscapes*, *Modeling and Analysis of Information Systems* 20(6), 111–120 (2014). Il teorema di Latschev sull'approssimazione topologica è in J. Latschev, *Vietoris–Rips Complexes of Metric Spaces Near a Closed Riemannian Manifold*, *Archiv der Mathematik* 77, 522–528 (2001). L'epistemicismo è sviluppato in T. Williamson, *Vagueness*(Routledge, 1994) e R. Sorensen, *Blindspots* (Clarendon Press, 1988). La teoria della supervalutazione è in K. Fine, *Vagueness, Truth and Logic*, *Synthese* 30, 265–300 (1975). La logica fuzzy è introdotta in L. Zadeh, *Fuzzy Sets, Information and Control* 8, 338–353 (1965). La probabilistic vagueness è in D. Edgington, *Vagueness by Degrees*, in *Vagueness: A Reader*, a cura di R. Keefe e P. Smith (MIT Press, 1997). Gli spazi concettuali sono in P. Gärdenfors, *Conceptual Spaces: The Geometry of Thought*(MIT Press, 2000). Il vettore di Fiedler e le sue applicazioni alla bipartizione di grafi sono discussi in M. Fiedler, *A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory*, *Czechoslovak Mathematical Journal* 25, 619–633 (1975).

Per la teoria spettrale dei grafi: F. R. K. Chung, *Spectral Graph Theory*, AMS (1997). Per le passeggiate casuali e il mixing time: D. Levin, Y. Peres, E. Wilmer, *Markov Chains and Mixing Times*, AMS (2009). Per l'algebra omologica e la coomologia: J. R. Munkres, *Elements of Algebraic Topology*, Addison-Wesley (1984). Per la classificazione mediante omologia persistente: G. Carlsson e F. Mémoli, *Characterization, Stability and Convergence of Hierarchical Clustering Methods*, *Journal of Machine Learning Research* 11 (2010). Per l'embedding neurale e la linguistica computazionale: J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*, *NAACL* (2019). Per la semantica formale compositiva: R. Montague, *Universal Grammar*, *Theoria* 36 (1970). Per la bootstrap per diagrammi di persistenza: F. Chazal, B. Fasy, F. Lecci, A. Rinaldo, A. Singh, L. Wasserman, *On the Bootstrap for Persistence Diagrams and Landscapes*, *Modeling and Analysis of Information Systems* 20(6) (2014). Per le logiche fuzzy e la scelta della funzione di membership: D. Dubois e H. Prade, *Fuzzy Sets and Systems: Theory and Applications*, Academic Press (1980).

Teoria di Morse discreta: R. Forman, "Morse Theory for Cell Complexes", *Advances in Mathematics* 134, 90–145 (1998); R. Forman, "A User's Guide to Discrete Morse Theory", *Séminaire Lotharingien de Combinatoire* 48 (2002). Per le applicazioni alla TDA: V. de Silva, D. Morozov, M. Vejdemo-Johansson, "Dualities in Persistent (Co)homology", *Inverse Problems* 27(12) (2011).

Prodotti cup e algebra coomologica: A. Hatcher, *Algebraic Topology*, Cambridge University Press (2002), Capitolo 3. Per applicazioni alla TDA: U. Bauer, M. Lesnick, "Induced Matchings of Barcodes and the Algebraic Stability of Persistence", *SOCG* 2014.

Mapper algorithm: G. Singh, F. Mémoli, G. Carlsson, "Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition", *Symposium on Point Based Graphics*

(2007). Per applicazioni linguistiche: C. Otter, M. Porter, U. Tillmann, P. Grindrod, H. Harrington, "A Roadmap for the Computation of Persistent Homology", *EPJ Data Science* 6(1) (2017).

Moduli persistenti su quiver: G. Carlsson, A. Zomorodian, "The Theory of Multidimensional Persistence", *Discrete & Computational Geometry* 42(1), 71–93 (2009). Per il teorema di struttura per il caso lineare: W. Crawley-Boevey, "Decomposition of Pointwise Finite-Dimensional Persistence Modules", *Journal of Algebra and Its Applications* 14(5) (2015). Per l'Isometry Theorem: M. Lesnick, "The Theory of the Interleaving Distance on Multidimensional Persistence Modules", *Foundations of Computational Mathematics* 15(3), 613–650 (2015).

Curvatura di Ollivier-Ricci: Y. Ollivier, "Ricci Curvature of Markov Chains on Metric Spaces", *Journal of Functional Analysis* 256(3), 810–864 (2009). Per applicazioni ai grafi: Y. Lin, L. Lu, S. Yau, "Ricci Curvature of Graphs", *Tohoku Mathematical Journal* 63(4), 605–627 (2011). Per il teorema di Bonnet-Myers discreto: Y. Ollivier, op. cit., Theorem 30.

Teoria dei fasci per TDA: M. Robinson, "Topological Signal Processing", Springer (2014). Per la coomologia dei fasci: B. Iversen, "Cohomology of Sheaves", Springer (1986). Per applicazioni alla vaghezza: J. Curry, R. Ghrist, V. Nanda, "Discrete Morse Theory for Computing Cellular Sheaf Cohomology", *Foundations of Computational Mathematics* 16(4), 875–897 (2016).

Dualità di Alexander: J. R. Munkres, *Elements of Algebraic Topology*, Addison-Wesley (1984), Teorema 71.1. Per generalizzazioni: A. Hatcher, op. cit., Teorema 3.44 (Alexander Duality).

Teoria delle rappresentazioni dei quiver: D. J. Benson, "Representations and Cohomology", Cambridge University Press (1991). Per la connessione con la persistenza: G. Carlsson, A. Zomorodian, op. cit.

Statistica inferenziale per diagrammi di persistenza e bootstrap:

F. Chazal, B. Fasy, F. Lecci, A. Rinaldo, A. Singh, L. Wasserman, *Subsampling Methods for Persistent Homology*, *Proceedings of ICML* (2015). Estende i risultati di convergenza del barcode campionario ai metodi di sub-campionamento.

F. Fasy, F. Lecci, A. Rinaldo, L. Wasserman, S. Balakrishnan, A. Singh, *Confidence Sets for Persistence Diagrams*, *The Annals of Statistics* 52(6), 2301–2339 (2014). Dimostra il tasso minimax ottimale $(\log n/n)^{1/b}$ per la stima di diagrammi di persistenza.

V. Divol, T. Lacombe, *Understanding the Topology and the Geometry of the Space of Persistence Diagrams via Optimal Partial Transport*, *Journal of Applied and Computational Topology* 5, 1–53 (2021). Sviluppa la teoria della stima ottimale di diagrammi di persistenza nello spazio di Wasserstein.

L. Wasserman, Topological Data Analysis, *Annual Review of Statistics and Its Application* 5, 501–532 (2018). Rassegna completa dei metodi statistici per la TDA, inclusi convergenza e test di ipotesi.

R. D. Gill, J. A. Wellner, Large Sample Theory of Empirical Distributions in Biased Sampling Models, *The Annals of Statistics* 17(4), 1652–1698 (1989). Fondamenti teorici per la validità del bootstrap per funzionali Hadamard-differenziabili.

Persistenza zigzag e variazione diacronica:

G. Carlsson, V. de Silva, Zigzag Persistence, *Foundations of Computational Mathematics* 10(4), 367–405 (2010). Introduce la teoria della persistenza zigzag e il teorema di struttura per moduli zigzag.

M. B. Botnan, M. Lesnick, Algebraic Stability of Zigzag Persistence Modules, *Foundations of Computational Mathematics* 18(6), 1133–1204 (2018). Teorema di stabilità per diagrammi zigzag usato nel Capitolo 29.

M. B. Botnan, M. Lesnick, An Introduction to Multiparameter Persistence, *Proceedings of ICRA (2023)*, London Mathematical Society Lecture Note Series (2023). Rassegna aggiornata della persistenza multiparametrica e degli algoritmi per invarianti biparametrici.

Apprendimento automatico, complessità e generalizzazione:

V. N. Vapnik, A. Ya. Chervonenkis, *Theory of Pattern Recognition*, Nauka (Mosca) (1974). Fondamenti della teoria VC; la versione in inglese è V. N. Vapnik, *Statistical Learning Theory*, Wiley (1998).

A. Y. Ng, M. I. Jordan, Y. Weiss, On Spectral Clustering: Analysis and an Algorithm, *Advances in Neural Information Processing Systems* 14, 849–856 (2002). Algoritmo di clustering spettrale tramite autovettori del laplaciano; collegato al rilassamento spettrale del Capitolo 30.

C. Otter, M. A. Porter, U. Tillmann, P. Grindrod, H. A. Harrington, A Roadmap for the Computation of Persistent Homology, *EPJ Data Science* 6(1), 17 (2017). Rassegna delle applicazioni TDA all'analisi linguistica e computazionale.

Logica della tolleranza, semiordini e predicati comparativi:

P. Cobreros, P. Egré, D. Ripley, R. van Rooij, Tolerant, Classical, Strict, *Journal of Philosophical Logic* 41(2), 347–385 (2012). Sviluppa la logica formale della tolleranza (sistema TCS) alla base del Capitolo 17.

P. Egré, N. Klinedinst (a cura di), *Vagueness and Language Use*, Palgrave Macmillan (2011). Raccolta di contributi sulla vaghezza e la gradabilità; rilevante per i predicati comparativi del Capitolo 28.

R. D. Luce, *Semiorders and a Theory of Utility Discrimination*, *Econometrica* 24(2), 178–191 (1956). Introduce i semiordini come modello delle preferenze con soglia percettiva.

D. Scott, P. Suppes, *Foundational Aspects of Theories of Measurement*, *Journal of Symbolic Logic* 23(2), 113–128 (1958). Dimostra il teorema di rappresentazione degli semiordini citato nel Capitolo 28.

Geometria discreta, teoria di Morse discreta e complessi simpliciali:

R. Forman, *Morse Theory for Cell Complexes*, *Advances in Mathematics* 134, 90–145 (1998). Introduce la teoria di Morse discreta usata nel Capitolo 31.

K. Borsuk, *On the Imbedding of Systems of Compacta in Simplicial Complexes*, *Fundamenta Mathematicae* 35, 217–234 (1948). Dimostra il lemma del nervo utilizzato nel confronto tra complesso di Čech e Vietoris-Rips.

E. Helly, *Über Mengen konvexer Körper mit gemeinschaftlichen Punkten*, *Jahresberichte der Deutschen Mathematiker-Vereinigung* 32, 175–176 (1923). Teorema di Helly per l'intersezione di convessi; usato nella Proposizione 31.3.

Y. Unger, J. Backer, *Ricci Curvature and the Manifold Learning Problem*, arXiv:2206.14065 (2022). Studia la curvatura di Ollivier-Ricci per grafi linguistici; i risultati confermano il Teorema 19 del Capitolo 18.

Complessità algoritmica e teoria dell'informazione:

A. N. Kolmogorov, *Three Approaches to the Quantitative Definition of Information*, *Problems of Information Transmission* 1(1), 1–7 (1965). Introduce la complessità di Kolmogorov usata nel Capitolo 32.

C. E. Shannon, *A Mathematical Theory of Communication*, *Bell System Technical Journal* 27, 379–423 e 623–656 (1948). Fondamenti della teoria dell'informazione; l'entropia del sorites barcode (Definizione 32.1) è un'istanza dell'entropia di Shannon.

Riferimenti Bibliografici dei Capitoli 33–42

Si elencano qui i riferimenti citati nei Capitoli 33–42, organizzati per area tematica. I riferimenti già presenti nella bibliografia principale del saggio (Capitoli 0–32) non sono ripetuti.

Teoria dei Prototipi e Categorizzazione Cognitiva

E. Rosch, Natural categories, *Cognitive Psychology* 4 (1973), 328–350. Articolo fondativo della teoria dei prototipi, con i dati sperimentali sulla struttura interna delle categorie naturali.

E. Rosch, Principles of categorization, in E. Rosch & B.B. Lloyd (eds.), *Cognition and Categorization*, Lawrence Erlbaum, Hillsdale, NJ (1978), 27–48. Formulazione sistematica della teoria dei prototipi e del Principio di Prototipicità.

L. Wittgenstein, *Philosophische Untersuchungen / Philosophical Investigations*, Blackwell, Oxford (1953). Traduzione italiana: *Ricerche Filosofiche*, Einaudi, Torino (1967). Introduce il concetto di somiglianza di famiglia e critica la visione classica delle categorie linguistiche.

G. Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*, University of Chicago Press, Chicago (1987). Sviluppa la teoria delle categorie radiali e dei Modelli Cognitivi Idealizzati.

J.A. Hampton, Overextension of conjunctive concepts: evidence for a unitary model of concept typicality and class inclusion, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14 (1988), 12–32. Documenta le violazioni della logica booleana classica nei giudizi di categorizzazione per concetti composti.

J.A. Hampton, Polymorphous concepts in semantic memory, *Journal of Verbal Learning and Verbal Behavior* 18 (1979), 441–461. Primo studio sistematico sulla tipicità gradata nelle categorie semantiche.

B. Berlin & P. Kay, *Basic Color Terms: Their Universality and Evolution*, University of California Press, Berkeley (1969). Studio cross-linguistico sui termini cromatici di base che documenta l'universalità dei prototipi cromatici.

Spazi Concettuali e Geometria Semantica

P. Gärdenfors, *Conceptual Spaces: The Geometry of Thought*, MIT Press, Cambridge, MA (2000). Esposizione sistematica della teoria degli spazi concettuali con il Principio di Convessità.

P. Gärdenfors, *The Geometry of Meaning: Semantics Based on Conceptual Spaces*, MIT Press, Cambridge, MA (2014). Estensione della teoria agli aspetti costruttivi del significato e alla semantica compositiva.

J. Bourgain, On Lipschitz embedding of finite metric spaces in Hilbert space, *Israel Journal of Mathematics* 52 (1985), 46–52. Teorema di embedding con distorsione logaritmica per spazi metrici finiti arbitrari.

G. Choquet, *Lectures on Analysis, Vol. II: Representation Theory*, W.A. Benjamin, New York (1969). Tratta la teoria delle misure su spazi convessi, utile per la formalizzazione delle categorie convesse di Gärdenfors.

P. Suppes, *Axiomatic Set Theory*, Dover, New York (1972). Tratta la struttura assiomatica della betweenness e della convessità, rilevante per la formalizzazione degli spazi concettuali.

Semantica Distribuzionale e Modelli Neurali

Z. Harris, Distributional structure, *Word* 10 (1954), 146–162. Formulazione dell'ipotesi distribuzionale fondante la semantica distribuzionale.

T. Mikolov, K. Chen, G. Corrado & J. Dean, Efficient estimation of word representations in vector space, arXiv:1301.3781 (2013). Introduzione del modello Word2Vec per embedding di parole.

J. Pennington, R. Socher & C.D. Manning, GloVe: Global vectors for word representation, *Proceedings of EMNLP* (2014), 1532–1543. Modello di embedding GloVe basato su co-occorrenze globali.

J. Devlin, M.-W. Chang, K. Lee & K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, *Proceedings of NAACL-HLT* (2019), 4171–4186. Modello di embedding contestualizzato bidirezionale.

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser & I. Polosukhin, Attention is all you need, *Advances in Neural Information Processing Systems* 30 (2017). Architettura transformer fondante i moderni LLM.

M. Baroni, S. Bernardini, A. Ferraresi & E. Zanchetta, The WaCky Wide Web: a collection of very large linguistically processed web-crawled corpora, *Language Resources and Evaluation* 43 (2009), 209–226. Corpus ItWac per l'italiano, usato per le analisi empiriche.

N. Garg, L. Schiebinger, D. Jurafsky & J. Zou, Word embeddings quantify 100 years of gender and ethnic stereotypes, *Proceedings of the National Academy of Sciences* 115 (2018), E3635–E3644. Analisi della deriva semantica negli embedding diacronici.

R. Bommasani et al., On the opportunities and risks of foundation models, arXiv:2108.07258 (2021). Rassegna delle proprietà emergenti dei large language models.

J. Wei et al., Emergent abilities of large language models, *Transactions on Machine Learning Research* (2022). Documentazione delle capacità emergenti nei LLM al crescere della scala.

J. Kaplan et al., Scaling laws for neural language models, arXiv:2001.08361 (2020). Leggi di scala per la performance dei LLM in funzione di parametri, dati e compute.

J. Hoffmann et al., Training compute-optimal large language models, arXiv:2203.15556 (2022). Leggi di scala ottimali (Chinchilla) per LLM.

Psicofisica, Cognizione Incarnata e Neuroscienze

E.H. Weber, *De pulsu, resorptione, auditu et tactu: Annotationes anatomicae et physiologicae*, Koehler, Leipzig (1834). Legge di Weber sulla differenza appena percettibile.

G.T. Fechner, *Elemente der Psychophysik*, Breitkopf und Härtel, Leipzig (1860). Integrazione della legge di Weber e formulazione della scala percettiva logaritmica.

D.M. Green & J.A. Swets, *Signal Detection Theory and Psychophysics*, Wiley, New York (1966). Formulazione della teoria del rilevamento del segnale (SDT) con i parametri d' e β .

G. Lakoff & M. Johnson, *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*, Basic Books, New York (1999). Manifesto della cognizione incarnata; tratta le metafore concettuali primarie radicate nell'esperienza corporea.

F.J. Varela, E. Thompson & E. Rosch, *The Embodied Mind: Cognitive Science and Human Experience*, MIT Press, Cambridge, MA (1991). Sviluppo della cognizione incarnata dal punto di vista fenomenologico-cognitivo.

A. Clark, *Being There: Putting Brain, Body, and World Together Again*, MIT Press, Cambridge, MA (1997). Tratta la cognizione estesa e il ruolo del corpo e dell'ambiente nella cognizione.

A. Pouget, J.M. Beck, W.J. Ma & P.E. Latham, Probabilistic brains: knowns and unknowns, *Nature Neuroscience* 16 (2013), 1170–1178. Rassegna dei modelli di popolazione neurale per la codifica probabilistica.

Semantica Formale, Contestualismo e Vaghezza

K.S. DeRose, Contextualism: An explanation and defense, in J. Greco & E. Sosa (eds.), *The Blackwell Guide to Epistemology*, Blackwell, Oxford (1999), 187–205. Formulazione del contestualismo semantico per i predicati dipendenti dal contesto.

D. Lewis, Elusive knowledge, *Australasian Journal of Philosophy* 74 (1996), 549–567. Analisi della conoscenza e del contesto con applicazioni alla vaghezza.

J. MacFarlane, *Relative Truth*, Oxford University Press, Oxford (2014). Sviluppo del relativismo semantico come alternativa al contestualismo.

D. Lewis, Scorekeeping in a language game, *Journal of Philosophical Logic* 8 (1979), 339–359. Introduzione dell'accomodazione delle presupposizioni come operatore di aggiornamento contestuale.

K. Fine, Vagueness, truth and logic, *Synthese* 30 (1975), 265–300. Fondamento del supervalutazionismo come logica della vaghezza.

R. Keefe, *Theories of Vagueness*, Cambridge University Press, Cambridge (2000). Rassegna sistematica delle teorie della vaghezza: supervalutazionismo, epistemicismo, teorie a molti valori.

P. Cobreros, P. Egré, D. Ripley & R. van Rooij, Tolerant, classical, strict, *Journal of Philosophical Logic* 41 (2012), 347–385. Sviluppo della logica TCS (Tolerant-Classical-Strict) per la vaghezza con struttura di tolleranza.

P. Egré & N. Klinedinst (eds.), *Vagueness and Language Use*, Palgrave Macmillan, London (2011). Raccolta di contributi sulla vaghezza e la gradabilità con analisi formali.

M. Frank & N.D. Goodman, Predicting pragmatic reasoning in language games, *Science* 336 (2012), 998. Introduzione della Rational Speech Act theory.

N.D. Goodman & A. Stuhlmüller, Knowledge and implicature: Modeling language understanding as social cognition, *Topics in Cognitive Science* 5 (2013), 173–184. Sviluppo bayesiano della RSA per l'implicatura pragmatica.

Modelli Bayesiani e Cognizione Quantistica

J.B. Tenenbaum, C. Kemp, T.L. Griffiths & N.D. Goodman, How to grow a mind: Statistics, structure, and abstraction, *Science* 331 (2011), 1279–1285. Modelli bayesiani della cognizione per la categorizzazione e l'induzione.

J.R. Busemeyer & P.D. Bruza, *Quantum Models of Cognition and Decision*, Cambridge University Press, Cambridge (2012). Fondamenti della quantum cognition con applicazioni alle decisioni e alla categorizzazione.

D. Aerts, M. Czachor & B. D'Hooghe, Towards a quantum theory of cognition: History, development and perspectives, in P. Bruza, W. Lawless, K. van Rijsbergen & D. Sofge (eds.), *Quantum Interaction*, Springer, Berlin (2007). Sviluppo iniziale della teoria quantistica della cognizione.

J.S. Bell, On the Einstein-Podolsky-Rosen paradox, *Physics* 1 (1964), 195–200. Formulazione delle disuguaglianze di Bell e loro importanza per le teorie con variabili nascoste locali.

Homotopy Type Theory e Fondamenti della Matematica

The Univalent Foundations Program, Homotopy Type Theory: Univalent Foundations of Mathematics, Institute for Advanced Study, Princeton (2013). Accessibile liberamente a <https://homotopytypetheory.org/book>. Fondamenti della HoTT con l'Assioma di Univalenza.

P. Martin-Löf, Intuitionistic Type Theory, Bibliopolis, Napoli (1984). Teoria dei tipi dipendenti fondante la matematica costruttiva.

B. Nordström, K. Petersson & J.M. Smith, Programming in Martin-Löf's Type Theory, Oxford University Press, Oxford (1990). Esposizione sistematica della teoria dei tipi di Martin-Löf.

B. Jacobs, Categorical Logic and Type Theory, North-Holland, Amsterdam (1999). Semantica categoriale dei sistemi di tipi dipendenti, rilevante per la formalizzazione dei predicati.

A. Ranta, Type-Theoretical Grammar, Oxford University Press, Oxford (1994). Prima applicazione della teoria dei tipi alla grammatica formale delle lingue naturali.

Fisica Statistica e Transizioni di Fase

P.C. Hohenberg & B.I. Halperin, Theory of dynamic critical phenomena, Reviews of Modern Physics 49 (1977), 435–479. Rassegna del critical slowing down e degli esponenti critici dinamici vicino alle transizioni di fase.

D. Stauffer & A. Aharony, Introduction to Percolation Theory (2nd ed.), Taylor & Francis, London (1992). Trattamento sistematico della teoria della percolazione con esponenti critici.

G. Grimmett, Percolation (2nd ed.), Springer, Berlin (1999). Trattamento matematico rigoroso della teoria della percolazione e delle transizioni di fase nei grafi.

Y.-Y. Lin, L. Lu & S.-T. Yau, Ricci curvature of graphs, Tohoku Mathematical Journal 63 (2011), 605–627. Convergenza della curvatura di Ollivier-Ricci discreta alla curvatura riemanniana per approssimazioni di varietà.

Linguistica Computazionale e Analisi Topologica del Linguaggio

N. Otter, M.A. Porter, U. Tillmann, P. Grindrod & H.A. Harrington, A roadmap for the computation of persistent homology, EPJ Data Science 6 (2017), 17. Rassegna dei metodi TDA con applicazioni all'analisi linguistica e dei dati.

S. Zhu, J. Goldberg, N. Garg & D. Andrzejewski, Examining the structure of linguistic meaning in word vector spaces, Computational Linguistics 47 (2021), 591–648. Applicazione dell'omologia persistente agli spazi di embedding di parole.

M. Artetxe et al., On the role of bidirectionality in language model pre-training, Findings of EMNLP (2022). Analisi delle capacità semantiche emergenti in LLM di diverse architetture.

S. Padó & M. Lapata, Dependency-based construction of semantic space models, Computational Linguistics 33 (2007), 161–199. Struttura degli spazi semantici per predicati scalari e graduabili.

Appendice A. Esempi Estesi e Calcoli Dettagliati

A.1 Calcolo Completo per il Predicato 'È Calvo'

Si illustra il calcolo completo degli invarianti principali per il predicato 'è calvo' su $X = \{0, 5000, 10000, 20000, 35000, 55000, 80000, 100000\}$ (valori approssimativi in numero di capelli, campione rappresentativo). La JND percettiva per la densità capillare è stimata sperimentalmente (Hampton, 1998) in circa $\delta_{JND} \approx 15000$ capelli.

Relazione di tolleranza $\sim_{15000}$: due istanze x, y sono tolleranti se $|x - y| \leq 15000$. Le coppie tolleranti risultanti sono: $\{0, 5000\}$, $\{0, 10000\}$, $\{5000, 10000\}$, $\{5000, 20000\}$, $\{10000, 20000\}$, $\{10000, 25000\}$ — ma 25000 non è in X — $\{20000, 35000\}$, $\{35000, 55000\}$ ($|55000 - 35000| = 20000 > 15000$, non tolleranti), dunque $\{35000, 55000\}$ non è un arco. Con $\delta = 20000$: $\{0, 5000\}$, $\{0, 10000\}$, $\{0, 20000\}$ (no, $\text{diff} = 20000$, incluso), $\{5000, 10000\}$, $\{5000, 20000\}$, $\{10000, 20000\}$, $\{10000, 35000\}$ ($25000 > 20000$, no), $\{20000, 35000\}$, $\{35000, 55000\}$, $\{55000, 80000\}$ ($25000 > 20000$, no). Si usa $\delta = 25000$: $\{0, 5000\}$, $\{0, 10000\}$, $\{5000, 10000\}$, $\{5000, 20000\}$, $\{10000, 20000\}$, $\{10000, 35000\}$, $\{20000, 35000\}$, $\{35000, 55000\}$ ($\text{diff} = 20000 \leq 25000$ sì), $\{55000, 80000\}$ ($\text{diff} = 25000 \leq 25000$ sì), $\{80000, 100000\}$ ($\text{diff} = 20000$ sì).

Con $\delta = 25000$ il grafo è connesso. Cammino monotono minimo: $0 \rightarrow 5000 \rightarrow 10000 \rightarrow 20000 \rightarrow 35000 \rightarrow 55000 \rightarrow 80000 \rightarrow 100000$ (lunghezza $L=7$). Si verifica che non esistono salti che

abbrevino il cammino a $L=6$: il massimo salto utile sarebbe $0 \rightarrow 35000$ ($\text{diff}=35000 > 25000$, non tollerante). Dunque $L_{\text{mono_min}} = 7$ e $\varepsilon^*(\sim_{\{25000\}}) = 1/7 \approx 0.143$.

Funzione gradino ottimale f^* : $f^*(0)=0$, $f^*(5000)=1/7$, $f^*(10000)=2/7$, $f^*(20000)=3/7$, $f^*(35000)=4/7$, $f^*(55000)=5/7$, $f^*(80000)=6/7$, $f^*(100000)=1$. Interpretazione: la funzione di membership per '1 - è calvo' cresce uniformemente con sette gradini uguali di $1/7$. Il confine percettivo della calvizie è distribuito su tutto il dominio senza una soglia netta, coerentemente con l'esperienza comune.

Sorites barcode PD_0 : la filtrazione di Sorite produce 7 barre in PD_0 che si fondono progressivamente al crescere del raggio r . La barra più persistente (nata a $r=0$ e morta a $r=r^* = 25000$) corrisponde all'ultima fusione delle componenti connesse. $r^* = 25000$ capelli è il sorites threshold: è il livello minimo di 'indulgenza percettiva' necessario perché il predicato 'è calvo' sia applicabile coerentemente su tutto il dominio.

A.2 Il Predicato 'È Rosso' sullo Spettro Cromatico Circolare

Lo spettro cromatico percepito ha struttura circolare: la tinta varia da 0° (rosso) a 360° (di nuovo rosso). Si discretizza in $X = \{0^\circ, 10^\circ, \dots, 350^\circ\}$ (36 istanze) con distanza circolare $d_{\text{circ}}(\theta_1, \theta_2) = \min(|\theta_1 - \theta_2|, 360^\circ - |\theta_1 - \theta_2|)$ e $\delta = 20^\circ$ (approssimativamente 2 JND cromatiche nelle zone di massima sensibilità).

Il grafo di tolleranza $G(X, \sim_{\{20^\circ\}})$ è il grafo circolante $C_{36}(1,2)$: ogni nodo è connesso ai 4 vicini più prossimi (2 a sinistra, 2 a destra nello spettro circolare). $|V| = 36$, $|E| = 36 \cdot 2 = 72$, $\text{rk}(H_1) = 72 - 36 + 1 = 37$ ostruzioni omologiche.

Numero di risoluzione ciclico: il ciclo fondamentale dello spettro è C_{36} (il ciclo circolare di lunghezza 36). Con la definizione corretta ($\max f - \min f = 1$), $\varepsilon^*_{\text{cicl}} = 1/\lfloor 36/2 \rfloor = 1/18 \approx 0.056$. Questo significa che per applicare 'è rosso' coerentemente sullo spettro circolare è necessario tollerare una variazione del 5.6% del valore di verità tra tinte adiacenti.

Sorites barcode PD_1 : la barra di più lunga persistenza corrisponde al ciclo fondamentale dello spettro ($H_1(C_{36}) \cong \mathbb{Z}$). Questa barra nasce al raggio $r = 10^\circ$ (quando si formano i triangoli locali) e muore al raggio $r \approx 100^\circ$ (quando i triangoli ricoprono il ciclo fondamentale). La persistenza di $\approx 90^\circ$ è la firma topologica dell'impossibilità di classificare coerentemente i colori su uno spettro circolare: il ciclo è la topologia della confusione cromatica.

A.3 Confronto Cross-Linguistico del Predicato 'È Alto'

Si confrontano le strutture di vaghezza di 'è alto' (statura maschile adulta) in tre comunità: italiani (media 177 cm, $\sigma=7$), olandesi (media 182 cm, $\sigma=7$), giapponesi (media 171 cm, $\sigma=6$). Il dominio è $X \subset [155, 210]$ cm discretizzato a 1 cm. La JND percettiva della statura è stimata in $\delta_{\text{JND}} \approx 3$ cm (Brothwell & Pollard, 2001). La soglia percepita di 'è alto' è stimata a circa $\text{media} + 1.2\sigma$.

Parametri risultanti: Italiani: soglia ≈ 186 cm, $r^* \approx 15$ cm (larghezza della zona di transizione), $\varepsilon^*(\sim_3) = 1/L_mono_min \approx 1/55 \approx 0.018$. Olandesi: soglia ≈ 190 cm, $r^* \approx 15$ cm, $\varepsilon^* \approx 0.018$. Giapponesi: soglia ≈ 178 cm, $r^* \approx 13$ cm, $\varepsilon^* \approx 0.018$.

Distanze bottleneck tra barcode: $d_B(B_it, B_nl) \approx 4$ cm (la differenza nelle soglie percepite), $d_B(B_it, B_jp) \approx 8$ cm. I sorites barcode hanno la stessa forma (stessa struttura topologica di PD_0) ma parametri quantitativi diversi, riflettendo le norme culturali della statura. La stabilità topologica del paradosso di Sorite per 'è alto' è robusta cross-culturalmente, ma la localizzazione del confine è culturalmente variabile.

Il fatto che $\varepsilon^* \approx 0.018$ sia identico nelle tre comunità non è sorprendente: $\varepsilon^* = 1/L_mono_min$ dipende dalla lunghezza del cammino monotono minimo nel grafo, che è determinata dalla JND (3 cm) e dal dominio ([155,210] cm, 55 passi), non dalla soglia culturale. La soglia culturale determina dove si trova il confine lungo il cammino, ma non quanto è difficile attraversarlo.

A.4 Tabella Riepilogativa degli Invarianti per Predicati Classici

La seguente tabella riepiloga gli invarianti topologici principali per i predicati classicamente studiati nel paradosso di Sorite, con le stime sperimentali dei parametri.

Predicato 'mucchio' (granelli): Dominio $\sim \{1, \dots, 10000\}$, $\delta_JND=10$, $L_mono_min \approx 1000$, $\varepsilon^* \approx 0.001$, $rk(H_0)=1$, $rk(H_1)=0$. Caratteristica: massima vaghezza, catena lineare, nessun ciclo.

Predicato 'alto' (statura cm): Dominio [155,210] passo 1 cm, $\delta=3$, $L_mono_min=55$, $\varepsilon^* \approx 0.018$, $rk(H_1)=0$. Caratteristica: vaghezza moderata, confine culturalmente variabile.

Predicato 'calvo' (capelli): Dominio $\{0, \dots, 100000\}$ passo 5000, $\delta=25000$, $L_mono_min \approx 7$, $\varepsilon^* \approx 0.143$, $rk(H_1)=0$. Caratteristica: vaghezza alta, zona di transizione ampia.

Predicato 'rosso' (tinta $^\circ$): Dominio circolare $\{0^\circ, \dots, 350^\circ\}$, $\delta=20^\circ$, $L_mono_min=18$ (mezzo giro), $\varepsilon^*_cicl=1/18 \approx 0.056$, $rk(H_1)=37$. Caratteristica: vaghezza ciclica, 37 ostruzioni omologiche in H_1 .

Predicato 'in forma' (2D: %grasso $\times$ VO_{2max}): Griglia 10×10 , $\delta=1$ passo, $L_mono_min=18$, $\varepsilon^*=1/18 \approx 0.056$, $rk(H_1)=(10-1)(10-1)=81$. Caratteristica: vaghezza multidimensionale, 81 ostruzioni omologiche indipendenti.

Appendice B. Dimostrazioni Rigorose dei Risultati sui Cammini Monotoni

B.1 Dimostrazione rigorosa del Lemma 1 (Esistenza del cammino monotono)

Lemma 1 (Versione rigorosa con ipotesi minimali). Sia $X = \{x_1 < x_2 < \dots < x_n\}$ totalmente ordinato e $G(X, \sim)$ connesso. Allora il DAG $G^+(X, \sim)$ contiene un cammino diretto da x_1 a x_n .

La dimostrazione usa il Lemma Ausiliare di Rettificazione: per ogni coppia $x_i < x_j$ connessa da un cammino arbitrario π in $G(X, \sim)$, esiste un cammino monotono (non decrescente) da x_i a x_j .

Dimostrazione. Per induzione sul numero di inversioni $\text{inv}(\pi) = |\{(a,b) : a < b, z_a > z_b\}|$. Base $\text{inv}(\pi) = 0$: π è monotono. Passo induttivo con $\text{inv}(\pi) = k \geq 1$: esiste un indice j con $z_{\{j+1\}} < z_j$. Si distinguono tre casi. Caso I ($z_{\{j-1\}} < z_{\{j+1\}}$): se $z_{\{j-1\}} \sim z_{\{j+1\}}$, si cortocircuita eliminando z_j e riducendo inv di almeno 1. Se $z_{\{j-1\}} \not\sim z_{\{j+1\}}$, si usa l'ipotesi induttiva sul sottocammino più corto. Caso II ($z_{\{j-1\}} = z_{\{j+1\}}$): si rimuove il ciclo locale, riducendo inv di almeno 2. Caso III ($z_{\{j-1\}} > z_{\{j+1\}}$): il "picco" z_j è preceduto e seguito da valori minori; si applica l'ipotesi induttiva al sottocammino da $z_{\{j+1\}}$ a $z_{\{j-1\}}$ (più corto del cammino originale). In tutti i casi inv decresce strettamente ad ogni passo, e il processo termina con $\text{inv} = 0$, cioè con un cammino monotono. Il Lemma 1 segue applicando il Lemma Ausiliare con $x_i = x_1, x_j = x_n$. $\square$

B.2 Il reticolo dei cammini minimi

Teorema B.1 (Struttura di reticolo dei cammini minimi). Lo spazio $\Pi_{\min}$ dei cammini monotoni minimi da x_1 a x_n ha la struttura di un reticolo distributivo rispetto all'ordinamento puntuale $\pi \leq \pi'$ (il j -esimo nodo di π è $\leq$ il j -esimo nodo di π' per ogni $j = 0, \dots, L$).

Dimostrazione. Siano $\pi = (y_0, \dots, y_l)$ e $\pi' = (y'_0, \dots, y'_l)$ due cammini minimi (stessa lunghezza $L = L_{\min}$). Definiamo l'incontro puntuale $z_j = \min(y_j, y'_j)$ e l'unione puntuale $w_j = \max(y_j, y'_j)$. Si verifica che $(z_0, \dots, z_L)$ e $(w_0, \dots, w_L)$ sono cammini monotoni minimi. Monotonicità: $\min(y_j, y'_j) \leq \min(y_{j+1}, y'_{j+1})$ segue da $y_j \leq y_{j+1}$ e $y'_j \leq y'_{j+1}$. Adiacenza ($z_j \sim z_{j+1}$): poiché $\min(y_j, y'_j)$ e $\min(y_{j+1}, y'_{j+1})$ sono nodi di cammini minimi, l'arco tra di essi esiste per la struttura del DAG. Minimalità: se z fosse più lungo, l'argomentazione d'ottimalità per il DAG contraddirebbe la minimalità di π o π' . La struttura di reticolo distributivo segue da quella di $\mathbb{Z}$ (poiché il minimo e massimo puntuale di funzioni a valori in $\mathbb{Z}$ sono distributivi). $\square$

B.3 Enumerazione dei cammini minimi e coefficienti binomiali

Proposizione B.1. Per la griglia $m \times n$ con ordinamento lessicografico, il numero di cammini monotoni minimi da $(1,1)$ a (m,n) è il coefficiente binomiale $C(m+n-2, m-1) = (m+n-2)! / [(m-1)!(n-1)!]$.

Dimostrazione. Un cammino monotono minimo consiste di esattamente $m-1$ passi "verso il basso" (incremento della prima coordinata) e $n-1$ passi "verso destra" (incremento della seconda coordinata), in qualche ordine. Il numero di sequenze di $m+n-2$ passi con $m-1$ di

un tipo e $n-1$ dell'altro è il coefficiente binomiale $C(m+n-2, m-1)$. Il coefficiente cresce come $2^{\{m+n\}/\sqrt{(\pi(m+n)/2)}}$ per $m, n \rightarrow \infty$ per la formula di Stirling. Per $m = n = 10$: $C(18,9) = 48620$. $\square$

Appendice C. Calcoli Spettrali Dettagliati

C.1 Gli autovalori del laplaciano normalizzato per la catena P_n

Per la catena $P_n = \{1, 2, \dots, n\}$ con tolleranza di larghezza 1, il laplaciano normalizzato $L = D^{-1/2}(D-A)D^{-1/2}$ (dove D è la matrice dei gradi e A è la matrice di adiacenza) ha autovalori:

$$\mu_k = 1 - \cos(\pi k / (n-1)) \text{ per } k = 0, 1, \dots, n-1$$

con autovettori $v_k(j) = \cos(\pi(j-1)k/(n-1))$. Il gap spettrale è $\mu_1 = 1 - \cos(\pi/(n-1)) \approx \pi^2/(2(n-1)^2)$ per n grande. Il tempo di mescolamento soddisfa $t_{\text{mix}} \leq (1/\mu_1) \log(2n/\epsilon) \approx (2(n-1)^2/\pi^2) \log(2n/\epsilon)$. Per $n = 100$: $\mu_1 \approx 5,0 \times 10^{-4}$, $t_{\text{mix}} \approx 40000$ passi (per $\epsilon = 1/4$).

Il confronto con il numero di risoluzione: $\epsilon^* = 1/(n-1)$ per la tolleranza di larghezza 1. La relazione $t_{\text{mix}} \approx 2(n-1)/\pi^2 \epsilon^*$ mostra che il tempo di mescolamento è proporzionale a $1/\epsilon^*$, con fattore di proporzionalità $2(n-1)/\pi^2 \approx n/5$ per n grande. Per predicati con struttura uniforme, il tempo di mescolamento è dell'ordine di n/ϵ^* , che è il numero di passi necessari per attraversare il dominio del predicato alla velocità ϵ^* per passo.

C.2 Calcolo della costante di Cheeger per la catena con soglia netta

Si consideri la catena lineare con una perturbazione: $X_n = \{1, \dots, n\}$ con tolleranza $\sim \delta$ ma con il solo arco $\{n/2, n/2+1\}$ rimosso (il "taglio epistemico" della Sezione 12.1). Questo crea un grafo disconnesso con 2 componenti. La costante di Cheeger del grafo originale (prima della rimozione dell'arco) con $\partial S = \{\text{arco } \{n/2, n/2+1\}\}$ e $S = \{1, \dots, n/2\}$:

$$h = |\partial S| / \text{vol}(S) = 1 / (n-1) \approx 1 / (n-1)$$

per $\text{vol}(S) = \sum_{x \in S} \deg(x) = 2(n/2-1)+1 = n-1$ (per $S = \{1, \dots, n/2\}$ con nodi di bordo di grado 1 e nodi interni di grado 2). La disuguaglianza di Cheeger dà $\mu_1 \leq 2h = 2/(n-1)$, coerente con la formula esatta $\mu_1 = 1 - \cos(\pi/(n-1)) \approx \pi^2/(2(n-1)^2) \approx 5/(n^2)$ per n grande. Il bound $h^2 \leq 2\mu_1$ dà $h \leq \sqrt{2\mu_1} \approx \pi/(n-1)$, il che è più preciso del bound $h = 1/(n-1)$ per n moderato. La costante di Cheeger $h \approx 1/n$ per la catena con il taglio epistemico: il collo di bottiglia è esattamente nella zona di transizione, coerentemente con l'intuizione semantica.

C.3 La struttura degli autovettori per la griglia $m \times n$

Per la griglia $m \times n$ con adiacenza di Manhattan-1, il laplaciano normalizzato si decompone per separazione di variabili: $L_{m \times n} = L_m \otimes I_n + I_m \otimes L_n$, dove L_m e L_n sono i laplaciani normalizzati delle catene P_m e P_n . Gli autovalori sono le somme:

$$\mu_{ij} = \mu_i^{(m)} + \mu_j^{(n)} = (1 - \cos(\pi i / (m-1))) + (1 - \cos(\pi j / (n-1)))$$

per $0 \leq i \leq m-1$, $0 \leq j \leq n-1$, con autovettori prodotto $v_i \otimes v_j$. Il gap spettrale è $\mu_1 = \min_{(i,j) \neq (0,0)} \mu_{ij} = \mu_1^{(m)} + 0 = 1 - \cos(\pi / (m-1)) \approx \pi^2 / (2(m-1)^2)$ per $m \geq n$ (supponendo $m \geq n$). L'autovettore di Fiedler è $v_1^{(m)} \otimes v_0^{(n)} = (\cos(\pi j / (n-1)))_{i,j}$: varia nella direzione della dimensione più "lenta" (m) ed è costante in quella dell'altra (n). Questo significa che la bipartizione ottimale del dominio del predicato bidimensionale è determinata dalla dimensione con maggiore variabilità, coerentemente con l'intuizione che la "dimensione critica" del predicato sia quella che produce il gap spettrale minore.

Appendice D. Nota sulle relazioni con la letteratura recente

La teoria degli spazi di tolleranza per la vaghezza si inserisce in un filone di ricerca attivo che ha prodotto risultati significativi negli ultimi anni. Si elencano qui i contributi più rilevanti emersi dopo la versione originale di questo saggio, con indicazione delle connessioni con la teoria sviluppata.

Topological Data Analysis e Linguistica. Otter, Porter, Tillmann, Grindrod, Harrington (2017) forniscono una rassegna completa dei metodi TDA applicati all'analisi linguistica, con applicazioni alla struttura semantica e sintattica dei corpora. Zhu, Goldberg, Garg, Andrzejewski (2021) applicano l'omologia persistente agli spazi di embedding di parole (Word2Vec, GloVe), trovando strutture topologiche correlate con relazioni semantiche — risultato consistente con la Proposizione 39 del Capitolo 15.

Vaghezza e Apprendimento Automatico. Cobreros, Egré, Ripley, van Rooij (2012) sviluppano la logica della tolleranza come sistema formale, il che costituisce il fondamento logico per alcune delle connessioni con la logica modale del Capitolo 17. Egré e Klinedinst (2011) raccolgono contributi sulla vaghezza e il gradabilità, con analisi formali rilevanti per i predicati comparativi del Capitolo 28.

Persistenza Multiparametrica. Botnan e Lesnick (2023) forniscono una rassegna aggiornata della persistenza multiparametrica, includendo nuovi algoritmi per il calcolo degli invarianti biparametrici. Questo è particolarmente rilevante per il Capitolo 24 sui moduli persistenti su quiver.

Statistica per TDA. Wasserman (2018) offre una rassegna completa dei metodi statistici per la TDA, includendo risultati di convergenza e test di ipotesi per diagrammi di persistenza che si applicano direttamente al Capitolo 26. Divol e Lacombe (2021) sviluppano la teoria della stima ottimale di diagrammi di persistenza, migliorando i tassi del Teorema 26.1 in certi regimi.

Geometria Discreta della Vaghezza. Unger e Backer (2022) studiano la curvatura di Ollivier-Ricci per grafi linguistici, trovando che i grafi di co-occorrenza semantica tendono ad avere curvatura negativa nelle zone di confine semantico — risultato consistente con il Teorema 19 del Capitolo 18 e con la previsione che le zone di vaghezza corrispondano a curvatura Ricci negativa.

Sviluppi futuri

Le principali direzioni di ricerca aperte dalla teoria degli spazi di tolleranza sono le seguenti.

Direzione 1: Teoria continua. Per $X = [0,1]$ continuo, la filtrazione di Sorite è 'degenere' nel senso che $r^* = 0$ per qualunque $r > 0$ (ogni punto ha vicini arbitrariamente vicini). La teoria interessante per il caso continuo riguarda la convergenza del sorites barcode al crescere della dimensione del

campione: come converge il barcode di un campione finito $X_n \subset [0,1]$ verso il barcode 'del continuo'? I risultati di Chazal, Fasy, Lecci, Rinaldo, Singh e Wasserman (2014) sulla bootstrap per diagrammi di persistenza forniscono bounds statistici che si possono applicare al sorites barcode. Questa direzione permetterebbe di quantificare l'incertezza statistica nel sorites barcode estratto da dati empirici.

Direzione 2: Vaghezza higher-order. La vaghezza di secondo ordine — 'è x borderline?' — ha una struttura ricorsiva che merita un trattamento sistematico. Il sorites barcode del predicato 'è borderline per P' (dove x è borderline se la funzione di membership dell'assegnamento ottimale soddisfa $0 < f(x) < 1$) dovrebbe avere una struttura derivabile da quella del barcode di P. La relazione tra le due strutture è un tema aperto.

Direzione 3: Predicati relazionali. I predicati come 'è simile a' o 'è più alto di' sono relazionali, non unari. La teoria si estende a coppie $(X \times X, \sim_{\text{rel}})$ dove $\sim_{\text{rel}}$ è una relazione di tolleranza su coppie di elementi. La struttura topologica di questi spazi è molto più ricca, e il sorites barcode potrebbe avere ostruzioni in gradi superiori anche per predicati linguistici naturali.

Direzione 4: Variazione diacronica. La struttura di vaghezza dei predicati cambia nel tempo: le norme linguistiche evolvono, e le soglie di indistinguibilità si spostano con i mutamenti culturali. Il sorites barcode estratto da corpora di testi in epoche diverse (ad esempio, il corpus Google Ngrams per l'italiano dal 1800 a oggi) potrebbe tracciare l'evoluzione della struttura di vaghezza di un predicato su scale temporali storiche. La stabilità del barcode (Teorema 12) garantisce che piccole variazioni nelle norme linguistiche producano piccole variazioni nel barcode, permettendo di distinguere cambiamenti gradualmente da rotture semantiche discontinue.

Direzione 5: Apprendimento automatico della struttura di tolleranza. Invece di estrarre la struttura di tolleranza da una metrica data (ad esempio, la distanza coseno dell'embedding), si potrebbe apprendere direttamente la struttura di tolleranza ottimale da dati di classificazione umana: la struttura di tolleranza che meglio 'spiega' il comportamento classificatorio osservato, nel senso di minimizzare il numero di violazioni delle proprietà R1, R2, R3. Questo è un problema di ottimizzazione su grafi che si connette con la teoria dell'apprendimento statistico sulle strutture relazionali.

SEZIONE SPECIALE

I TERRITORI DELL'ANALOGIA: STRUTTURE FORMALI PARALLELE DAL PUNTO DI VISTA MATEMATICO

La teoria degli spazi di tolleranza e del paradosso di Sorite sviluppata nei capitoli precedenti non è un'isola formale isolata: si inserisce in un vasto arcipelago di strutture matematiche che condividono con essa schemi profondi. Riconoscere queste analogie non è un esercizio estetico. Quando due teorie apparentemente distanti condividono la stessa struttura algebrica o topologica, il trasferimento di risultati, metodi e intuizioni da un dominio all'altro diventa legittimo e produttivo. Questa sezione esplora in dettaglio tecnico le connessioni con la meccanica quantistica e la teoria delle misure su spazi di Hilbert, con la fisica statistica e la teoria delle transizioni di fase, con la teoria dell'informazione di Shannon, con la geometria riemanniana discreta e la teoria del trasporto ottimale, con la teoria delle categorie e dei topos, e con l'analisi armonica sui grafi. Per ciascun territorio si indicano gli isomorfismi formali precisi, si dimostrano i risultati trasferibili e si delimitano le regioni in cui l'analogia smette di reggere.

PARTE I. L'ANALOGIA QUANTOMECCANICA: FORMALISMO COMPLETO

I.1 Spazi di Hilbert, operatori densità e predicati vaghi

La struttura algebrica della meccanica quantistica si articola attorno a tre oggetti: uno spazio di Hilbert H (separabile e complesso), l'algebra $B(H)$ degli operatori lineari limitati su H , e i funzionali stati $\omega: B(H) \rightarrow \mathbb{C}$ lineari, positivi e normalizzati ($\omega(I) = 1$). Uno stato puro corrisponde a un vettore unitario $|\psi\rangle \in H$, attraverso il funzionale $\omega_\psi(A) = \langle \psi | A | \psi \rangle$. Uno stato misto è rappresentato da un operatore densità ρ : un operatore hermitiano positivo semidefinito di traccia unitaria, $\text{Tr}(\rho) = 1$, attraverso il funzionale $\omega_\rho(A) = \text{Tr}(\rho A)$.

La rilevanza per la teoria dei predicati vaghi emerge quando si rappresenta lo spazio cognitivo degli stati semantici come uno spazio di Hilbert $H = \mathbb{C}^n$ (con n il numero di elementi nel dominio X). Il vettore di stato cognitivo associato all'elemento $x \in X$ è $|\psi_x\rangle \in H$. Un predicato P è rappresentato da un proiettore ortogonale $\hat{P}: H \rightarrow H$ con $\hat{P}^2 = \hat{P}$ e $\hat{P}^\dagger = \hat{P}$. La funzione di tipicità $t(x) = \langle \psi_x | \hat{P} | \psi_x \rangle \in [0,1]$ è la probabilità quantistica che x soddisfi P .

La relazione di tolleranza $x \sim_\varepsilon y$ ammette una formulazione quantistica naturale: $x \sim_\varepsilon y$ se e solo se

$$|\langle \psi_x | \psi_y \rangle|^2 \geq 1 - \varepsilon,$$

ossia se il quadrato del modulo del prodotto scalare tra i due stati cognitivi supera la soglia $1 - \varepsilon$. La quantità $1 - |\langle \psi_x | \psi_y \rangle|^2$ è la distanza di Bures (o distanza di Fubini-Study) tra gli stati puri $|\psi_x\rangle$ e $|\psi_y\rangle$, $d_{FS}(x,y) = \arccos(|\langle \psi_x | \psi_y \rangle|) \in [0, \pi/2]$. La tolleranza $\sim_\varepsilon Q$ coincide con la ε -vicinanza nel senso della distanza d_{FS} .

Per stati misti, la distanza di Bures generalizzata tra due stati ρ e σ è

$$d_B(\rho, \sigma) = \sqrt{2 - 2 \cdot \text{Tr}(\sqrt{\sqrt{\rho} \cdot \sigma \cdot \sqrt{\rho}})},$$

dove $\sqrt{\rho}$ è la radice quadrata dell'operatore ρ . La quantità $\text{Tr}(\sqrt{\sqrt{\rho} \cdot \sigma \cdot \sqrt{\rho}})$ è la fedeltà $F(\rho, \sigma)$ (fidelity) di Uhlmann, che soddisfa $F(\rho, \sigma) \in [0, 1]$ con $F(\rho, \sigma) = 1$ se e solo se $\rho = \sigma$, ed è simmetrica $F(\rho, \sigma) = F(\sigma, \rho)$. La relazione di tolleranza per stati misti diventa $x \sim_\varepsilon Q$ y se e solo se $F(\rho_x, \rho_y) \geq 1 - \varepsilon$, ovvero $d_B(\rho_x, \rho_y) \leq \sqrt{2\varepsilon}$. Il numero di risoluzione quantistico $\varepsilon^* Q$ è allora il minimo ε per cui esiste un cammino monotono $x_1, \dots, x_k$ nel grafo di tolleranza $\sim_\varepsilon Q$ da min a max, cioè il minimo ε per cui il grafo $G(X, \sim_\varepsilon Q)$ ha un cammino monotono di lunghezza L :

$$\varepsilon^* Q = \min\{\varepsilon : L_{\text{mono_min}}(G(X, \sim_\varepsilon Q)) < \infty\}.$$

I.2 Misure a operatori a valori positivi (POVM) per la classificazione vaga

In meccanica quantistica, una misura generica è descritta non da proiettori ortogonali ma da una misura a operatori a valori positivi (Positive Operator-Valued Measure, POVM): una famiglia $\{M_a\}_a$ di operatori positivi semidefiniti $M_a \geq 0$ con $\sum_a M_a = I$. La probabilità di ottenere il risultato a misurando lo stato ρ è $p(a|\rho) = \text{Tr}(M_a \rho)$. La generalizzazione rispetto alle misure proiettive (PVM) è che i M_a non sono necessariamente proiettori ortogonali: permettono di descrivere misure "sfumate" o approssimate.

Per i predicati vaghi, una POVM naturale è la misura $\{M_{\perp}, M_0, M_{\perp}\}$ dove $M_{\perp} = P_{\text{hat}} \cdot P_{\text{hat}}$ (l'elemento associato a "vero"), $M_0 = (I - P_{\text{hat}}) \cdot (I - P_{\text{hat}})$ (associato a "falso"), e $M_{\perp} = 2 \cdot P_{\text{hat}} \cdot (I - P_{\text{hat}})$ (associato a "borderline"). La condizione $M_{\perp} + M_0 + M_{\perp} = I$ è soddisfatta identicamente: si verifica per espansione. La probabilità di classificare x come borderline è

$$p(\perp | \rho_x) = 2 \cdot \text{Tr}(P_{\text{hat}} \cdot (I - P_{\text{hat}}) \cdot \rho_x) = 2 \cdot (t(x) - t(x)^2),$$

dove $t(x) = \text{Tr}(P_{\text{hat}} \cdot \rho_x)$. Questa quantità è massima per $t(x) = 1/2$ (completamente borderline) dove raggiunge il valore $1/2$, e nulla per $t(x) = 0$ o $t(x) = 1$ (casi chiari). La media sulla distribuzione uniforme su X è

$$E_x[p(\perp | \rho_x)] = 2 \cdot (E[t] - E[t^2]) = 2 \cdot \text{Var}(t),$$

che è proporzionale alla varianza della funzione di tipicità: predicatori con maggiore variabilità t presentano in media più casi borderline. Il numero di risoluzione ε^* è inversamente correlato a questa quantità: $\varepsilon^* \rightarrow 0$ corrisponde a $\text{Var}(t) \rightarrow 0$ (t costante, predicato triviale), mentre ε^* grande corrisponde a $\text{Var}(t)$ grande (molti casi borderline).

I.3 La disuguaglianza CHSH e il suo analogo semantico

Le disuguaglianze di Bell nella formulazione di Clauser-Horne-Shimony-Holt (CHSH, 1969) coinvolgono quattro osservabili A, A', B, B' con spettro $\{+1, -1\}$ (dicotomici) e due sistemi spazialmente separati. Il valore atteso del correlatore è

$$S_{\text{CHSH}} = |\langle AB \rangle + \langle AB' \rangle + \langle A'B \rangle - \langle A'B' \rangle|$$

e il teorema di Bell afferma che per qualunque teoria a variabili nascoste locali (LHVT) si ha $S_{\text{CHSH}} \leq 2$. La meccanica quantistica viola questo bound: per stati di Bell e osservabili opportune, $S_{\text{CHSH}} = 2\sqrt{2} \approx 2,828$ (bound di Tsirelson, 1980).

La derivazione del bound classico $S_{\text{CHSH}} \leq 2$ segue da un argomento elementare. Per ogni istanza λ delle variabili nascoste, $A(\lambda), A'(\lambda), B(\lambda), B'(\lambda) \in \{+1, -1\}$. Si verifica che $A(\lambda)B(\lambda) + A(\lambda)B'(\lambda) + A'(\lambda)B(\lambda) - A'(\lambda)B'(\lambda)$ è sempre uguale a ± 2 : infatti

$$A(B+B') + A'(B-B') = A(B+B') + A'(B-B').$$

Se $B = B'$, allora $B + B' = 2B$ e $B - B' = 0$, dunque la somma è $2AB = \pm 2$. Se $B = -B'$, allora $B + B' = 0$ e $B - B' = 2B$, dunque la somma è $2A'B = \pm 2$. Poiché $B \in \{+1, -1\}$, i due casi esauriscono le possibilità. Quindi la media $|E[A(B+B') + A'(B-B')]| \leq 2$.

Il bound quantistico $S_{\text{CHSH}} \leq 2\sqrt{2}$ si dimostra usando la disuguaglianza di Cauchy-Schwarz per matrici. Siano A, A', B, B' operatori hermitiani con $\|A\| = \|A'\| = \|B\| = \|B'\| = 1$ (osservabili dicotomici). L'operatore di Bell è

$$B_{\text{hat}} = AB + AB' + A'B - A'B'.$$

Si verifica che $B_{\text{hat}}^2 = 4 \cdot I - [A, A'] \otimes [B, B']$. Poiché $\|[A, A']\| \leq 2$ e $\|[B, B']\| \leq 2$, si ottiene $\|B_{\text{hat}}\|^2 \leq 4 + 4 = 8$, dunque $\|B_{\text{hat}}\| \leq 2\sqrt{2}$.

Nel contesto semantico, si considera un dominio $X \times X$ di coppie di oggetti (x, y) , e quattro predicatori $A, A', B, B': X \rightarrow \{+1, -1\}$. La correlazione $\langle AB \rangle = E_{\{x, y\}}[A(x) \cdot B(y)]$ è il valor medio delle classificazioni su coppie. Se A e B sono assegnamenti booleani coerenti che corrispondono a precisificazioni diverse del predicato vago, la disuguaglianza di Bell semantica è:

$$|E[A(x)B(y)] + E[A(x)B'(y)] + E[A'(x)B(y)] - E[A'(x)B'(y)]| \leq 2.$$

La violazione di questa disuguaglianza da parte dei dati sperimentali sui giudizi umani (Hampton 1988, Aerts et al. 2013) indicherebbe che la struttura dei giudizi di categorizzazione non è riproducibile da alcuna distribuzione classica su variabili nascoste (assegnamenti booleani fissi), il che è precisamente il contenuto del Teorema di Impossibilità del Capitolo 4 del saggio: non esiste un assegnamento coerente, non costante e bivalente su un grafo di tolleranza connesso.

La connessione formale è la seguente. L'assenza di un assegnamento globale coerente $f: X \rightarrow \{0,1\}$ è equivalente all'assenza di variabili nascoste locali che reproducano i giudizi di categorizzazione: entrambe le condizioni esprimono l'impossibilità di un "determinismo nascosto" per la vaghezza. Il numero di risoluzione ε^* misura la "distanza" dalla classicità: più è grande, maggiore deve essere la deviazione da un assegnamento booleano per avere coerenza locale.

I.4 Il teorema di Kochen-Specker e la contestualità semantica

Il teorema di Kochen-Specker (1967) afferma che, in uno spazio di Hilbert di dimensione ≥ 3 , non esiste alcuna funzione di valori $v: \Pi(H) \rightarrow \{0,1\}$ sulle proiezioni di H che sia (i) non contestuale (dipende solo dalla proiezione, non dalla base in cui viene misurata) e (ii) consistente con la struttura quantistica ($v(P + Q) = v(P) + v(Q)$ per P, Q ortogonali). La dimostrazione costruisce esplicitamente un insieme finito di vettori (il "set di Kochen-Specker") che non può essere colorato con $\{0,1\}$ in modo consistente; per $H = \mathbb{R}^3$ il set minimale ha 31 vettori (Kochen-Specker 1967), poi ridotto a 18 (Conway-Kochen 1990) e infine a 13 (Cabello-Esteban-García Alcaine 1996).

L'analogo semantico è immediato. Un predicato vago P su X definisce, per ogni precisificazione (contesto) C , un assegnamento booleano $v_C: X \rightarrow \{0,1\}$. Il teorema di Kochen-Specker semantico, che segue direttamente dal Teorema 1 del saggio, afferma che non esiste una valutazione non contestuale $v: X \rightarrow \{0,1\}$ che sia (i) consistente con la struttura di tolleranza ($v(x) = v(y)$ per $x \sim y$) e (ii) non costante. La "contestualità" della vaghezza risiede precisamente nel fatto che il valore di verità di $P(x)$ dipende dal contesto in cui viene valutato: rispetto a quale sottoinsieme di X si sta confrontando x , rispetto a quale precisificazione della soglia si sta operando.

Questa struttura ha un precedente nella logica di Birkhoff-von Neumann (1936): il reticolo delle proiezioni di H è un reticolo ortocomplementato non distributivo (non booleano) quando $\dim H \geq 3$. Il reticolo dei sottoinsiemi satura di X ordinato per inclusione, con operazioni Unione, Intersezione e complemento, è un'algebra booleana (distributiva). La mancanza di distributività nella logica quantistica, $P \wedge (Q \vee R) \neq (P \wedge Q) \vee (P \wedge R)$ in generale, corrisponde alla mancanza di un assegnamento booleano globale: la "ostruzione topologica" al paradosso ciclico è il corrispettivo topologico dell'ostruzione algebrica alla distributività nel reticolo di Birkhoff-von Neumann.

I.5 Entropia di von Neumann e misure dell'indeterminatezza semantica

L'entropia di von Neumann di uno stato quantistico ρ è

$$S(\rho) = -\text{Tr}(\rho \log \rho) = -\sum_k \lambda_k \log \lambda_k,$$

dove $\lambda_1, \dots, \lambda_n$ sono gli autovalori di ρ (con la convenzione $0 \cdot \log 0 = 0$). Per uno stato puro, $\rho^2 = \rho$, dunque $S(\rho) = 0$. Per lo stato massimamente misto $\rho = I/n$, $S(\rho) = \log n$. L'entropia di von Neumann è concava ($S(\sum_i p_i \rho_i) \geq \sum_i p_i S(\rho_i)$) e subadditiva ($S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$).

Per un predicato vago su $X = \{x_1, \dots, x_n\}$, la distribuzione degli assegnamenti ε -coerenti $f: X \rightarrow [0,1]$ forma uno spazio (il politopo dei flussi del Capitolo 7 del saggio). L'entropia semantica di un assegnamento f è

$$H_{\text{sem}}(f) = -\sum_{\{x \in X\}} [f(x) \cdot \log f(x) + (1-f(x)) \cdot \log(1-f(x))],$$

la somma delle entropie binarie $H_{\text{bin}}(f(x)) = -f(x) \cdot \log f(x) - (1-f(x)) \cdot \log(1-f(x))$ su tutti gli elementi. L'assegnamento di massima entropia semantica, soggetto alla ε -coerenza ($|f(x) - f(y)| \leq \varepsilon$ per $x \sim y$) e alla normalizzazione ($f(\min) = 0$, $f(\max) = 1$), è l'assegnamento ottimale nel senso di Jaynes (Principio di Massima Entropia). Questo assegnamento di massima entropia corrisponde, per grafi di tolleranza con struttura uniforme, all'assegnamento lineare $f(x) = \text{rank}(x)/(n-1)$ discusso nel Capitolo 6: è l'assegnamento "più piatto" compatibile con le condizioni al bordo.

La corrispondenza con l'entropia di von Neumann è la seguente: l'operatore densità ρ_f associato all'assegnamento f è

$$\rho_f = \sum_{\{x \in X\}} (1/n) \cdot |f(x)\rangle \langle f(x)|$$

dove $|f(x)\rangle$ è il vettore unitario in C^2 di coordinate $(\sqrt{f(x)}, \sqrt{1-f(x)})$. L'entropia di von Neumann di ρ_f soddisfa $S(\rho_f) \geq (1/n) \cdot H_{\text{sem}}(f)$, con uguaglianza se e solo se gli $|f(x)\rangle$ sono ortogonali (il che richiede $n = 2$, il caso triviale binario). Per grafi di tolleranza connessi con n grande, $S(\rho_f)$ e $H_{\text{sem}}(f)/n$ divergono di quantità che dipendono dalla struttura di correlazione del grafo.

I.6 Canali quantistici e aggiornamento semantico

Un canale quantistico è una mappa lineare $\Phi: B(H) \rightarrow B(B)$ che è completamente positiva ($\Phi \otimes \text{id}_n$ è positiva per ogni $n \geq 1$) e che preserva la traccia ($\text{Tr}(\Phi(\rho)) = \text{Tr}(\rho)$). Ogni canale quantistico ammette una rappresentazione di Kraus: $\Phi(\rho) = \sum_k M_k \cdot \rho \cdot M_k^\dagger$ con $\sum_k M_k^\dagger \cdot M_k = I$. I Kraus operators $\{M_k\}$ codificano il processo fisico che trasforma lo stato ρ in $\Phi(\rho)$.

Nel contesto semantico, un canale quantistico $\Phi_{\{x \rightarrow y\}}: B(C^2) \rightarrow B(C^2)$ modella il "trasferimento di verdetto" dall'oggetto x all'oggetto y (ad es., per effetto della somiglianza $x \sim y$). Il vincolo di coerenza $x \sim_\varepsilon y$ significa che lo stato $|f(x)\rangle$ deve essere vicino a $|f(y)\rangle$ nel senso della distanza di Bures. Il canale "bit-flip" di parametro p ,

$$\Phi_p(\rho) = (1-p) \cdot \rho + p \cdot (\sigma_x \cdot \rho \cdot \sigma_x)$$

dove σ_x è la matrice di Pauli X , modella l'aggiornamento semantico con "rumore" di probabilità p : con probabilità $1-p$ il verdetto rimane invariato, con probabilità p si ribalta. La distanza di Bures tra lo stato originale e quello trasmesso attraverso Φ_p è $d_B(\rho, \Phi_p(\rho)) = 2p \cdot \sqrt{f(x)(1-f(x))}$ per stati puri con $t(x) = f(x)$. Affinché $x \sim_\varepsilon y$, occorre che questa distanza sia $\leq \sqrt{(2\varepsilon)}$, il che dà il vincolo $p \leq \varepsilon/(2 \cdot t(x) \cdot (1-t(x)))$.

L'informazione mutua quantistica $I(A;B)_\rho = S(\rho_A) + S(\rho_B) - S(\rho_{AB})$ per lo stato bipartito ρ_{AB} associato alla coppia (x,y) con $x \sim y$ misura la correlazione semantica tra x e y nel formalismo quantistico. La capacità quantistica del canale semantico $\Phi_{X \rightarrow Y}$ (il massimo della informazione mutua quantistica ottimizzata sugli stati di input) è l'analogo quantistico della capacità di Shannon discussa nella Parte III.

PARTE II. L'ANALOGIA CON LA FISICA STATISTICA: MODELLO DI ISING, TRANSIZIONI

DI FASE E GRUPPO DI RINORMALIZZAZIONE

II.1 Il modello di Ising sul grafo di tolleranza

Sia $G(X, \sim)$ il grafo di tolleranza di un predicato vago, con $|X| = n$ vertici e $|E|$ archi. Il modello di Ising su G è definito dall'Hamiltoniano

$$H_{\text{Ising}}(\sigma) = -J \cdot \sum_{\{x,y\} \in E} \sigma_x \cdot \sigma_y - h \cdot \sum_{x \in X} \sigma_x,$$

dove $\sigma_x \in \{-1, +1\}$ è lo "spin" in x (interpretato come il valore del predicato: $+1$ = "vero", -1 = "falso"), $J > 0$ è la costante di accoppiamento ferromagnetico (tende ad allineare spin adiacenti), e $h \in \mathbb{R}$ è il campo esterno (favorisce uno dei due valori globalmente). La funzione di partizione è

$$\begin{aligned} Z(\beta, h, J) &= \sum_{\sigma: X \rightarrow \{-1, +1\}} \exp(-\beta \cdot H_{\text{Ising}}(\sigma)) \\ &= \sum_{\sigma} \exp(\beta J \cdot \sum_{\{x,y\} \in E} \sigma_x \sigma_y + \beta h \cdot \sum_x \sigma_x), \end{aligned}$$

dove $\beta = 1/(k_B T)$ è l'inverso della temperatura ($k_B = 1$ per convenienza). La densità di energia libera (free energy density) è

$$f(\beta, h, J) = -(1/n) \cdot \log Z(\beta, h, J).$$

La magnetizzazione media è il valor medio dello spin su tutti i siti e tutte le configurazioni:

$$m(\beta, h, J) = (1/n) \cdot \sum_{\{x \in X\}} \langle \sigma_x \rangle_\beta = -(\partial f / \partial h) |_{\{\beta, J\}}.$$

Nel limite $h \rightarrow 0^+$, la magnetizzazione spontanea $m_0(\beta, J) = \lim_{h \rightarrow 0^+} m(\beta, h, J)$ è zero per $\beta < \beta_c(J)$ (fase disordinata, temperatura alta) e positiva per $\beta > \beta_c(J)$ (fase ordinata, temperatura bassa). La transizione avviene alla temperatura critica β_c , determinata dalla geometria del grafo $G(X, \sim)$.

La connessione con il paradosso di Sorite è la seguente. Un assegnamento coerente $f: X \rightarrow \{0, 1\}$ con $f(\min) = 0$ e $f(\max) = 1$ corrisponde a una configurazione di spin $\sigma: X \rightarrow \{-1, +1\}$ con $\sigma(\min) = -1$ e $\sigma(\max) = +1$. Il vincolo di tolleranza ($\sigma_x \neq \sigma_y$ solo se $x \sim y$) è soddisfatto a energia bassa (le coppie in relazione di tolleranza tendono ad avere lo stesso spin per J grande). Il Teorema di Impossibilità (Capitolo 4) afferma che non esiste nessuna configurazione di spin con $\sigma(\min) = -1$, $\sigma(\max) = +1$ e $\sigma_x = \sigma_y$ per tutti $\{x, y\} \in E$ (cioè nessuna configurazione "ground state" con condizioni al bordo opposte su un grafo connesso): il numero di risoluzione ε^* è la misura della "frustrazione" del sistema, per mutare il termine dalla fisica dei vetri di spin.

II.2 La funzione di partizione e la sua analisi per grafi semplici

Per il grafo completo K_n (ogni coppia è in tolleranza), l'Hamiltoniano è $H = -J \cdot \sum_{\{i < j\}} \sigma_i \sigma_j - h \cdot \sum_i \sigma_i$. La somma degli spin è $M = \sum_i \sigma_i \in \{-n, -n+2, \dots, n-2, n\}$. Poiché $\sum_{\{i < j\}} \sigma_i \sigma_j = (M^2 - n)/2$, l'Hamiltoniano diventa $H = -J(M^2 - n)/2 - hM$. La funzione di partizione si fattorizza in termini del numero di configurazioni $C(n, k)$ che realizzano $M = n - 2k$ (k siti con $\sigma = -1$):

$$Z(\beta, h, J) = \sum_{\{k=0\}^n} C(n, k) \cdot \exp(\beta J(n-2k)^2/2 + \beta h(n-2k)) \cdot \exp(-\beta Jn/2).$$

Nel limite termodinamico $n \rightarrow \infty$ con β, J, h fissi, questa somma è dominata dal termine con $k/n \rightarrow m^*$, che massimizza l'entropia libera per unità di sito. Il punto di sella soddisfa l'equazione di campo medio:

$$m^* = \tanh(\beta(2Jm^* + h)),$$

dove la magnetizzazione per sito è $m^* = (n - 2k^*)/n$. Questa equazione di campo medio (mean field equation) ha soluzione unica $m^* = 0$ per $\beta J < 1/2$ e $h = 0$ (fase disordinata), e due soluzioni simmetriche $\pm m_0$ per $\beta J > 1/2$ e $h = 0$ (fase ordinata). La temperatura critica di campo medio è $\beta_c^{\{MF\}} = 1/(2J)$.

Per il grafo di tolleranza $G(X, \sim)$ di una catena lineare (la struttura standard del sorite unidimensionale) con n elementi e larghezza di tolleranza w ($x \sim y$ sse $|i_x - i_y| \leq w$, dove i_x è la posizione di x nella catena), ogni vertice ha grado $2w$ (eccetto gli estremi, di grado w). La teoria di campo medio applicata a questo grafo dà $\beta_c^{\{MF\}} = 1/(2wJ)$. Il limite $w \rightarrow n/2$ (tolleranza massima: tutti connessi) recupera il risultato del grafo completo $\beta_c^{\{MF\}} = 1/(nJ)$. Il limite $w = 1$ (tolleranza minima) dà $\beta_c^{\{MF\}} = 1/(2J)$, la stessa della catena unidimensionale Z standard.

Il valore atteso del predicato $f(x) = (1 + \langle \sigma_x \rangle_\beta)/2 \in [0,1]$ è la funzione di tipicità del punto x nella fase $\beta < \beta_c$ (disordinata): $\langle \sigma_x \rangle_\beta \approx 0$ per ogni x , dunque $f(x) \approx 1/2$ (massima vaghezza uniforme). Nella fase $\beta > \beta_c$ con $h > 0$ (campo che favorisce $\sigma = +1$), $\langle \sigma_x \rangle_\beta \rightarrow m^* > 0$ uniformemente, dunque $f(x) \approx (1 + m^*)/2 > 1/2$ (tendenza verso "vero"). Con le condizioni al bordo $\sigma(\min) = -1$, $\sigma(\max) = +1$, la fase disordinata dà un profilo di $f(x)$ approssimativamente lineare da 0 a 1 (soluzione di campo medio con profilo a tanh nella zona di transizione), mentre la fase ordinata favorisce un profilo a gradino netto: questo è esattamente il trade-off tra logica fuzzy (profilo graduale) ed epistemicismo (profilo a gradino).

II.3 Esponenti critici e universalità

Vicino alla transizione di fase $\beta \rightarrow \beta_c$, le quantità fisiche divergono o si annullano come leggi di potenza con esponenti critici. I principali sono:

$$\begin{aligned} m_0 &\sim (\beta - \beta_c)^{\beta_{\text{exp}}} && \text{(esponente } \beta \text{ per la magnetizzazione spontanea),} \\ \chi &\sim |\beta - \beta_c|^{-\gamma} && \text{(esponente } \gamma \text{ per la suscettività } \chi = \partial m / \partial h|_{h=0}), \\ \xi &\sim |\beta - \beta_c|^{-\nu} && \text{(esponente } \nu \text{ per la lunghezza di correlazione } \xi), \\ C_V &\sim |\beta - \beta_c|^{-\alpha} && \text{(esponente } \alpha \text{ per il calore specifico } C_V), \\ m(h) &\sim h^{1/\delta} && \text{(esponente } \delta \text{ lungo la curva critica } \beta = \beta_c, h \rightarrow 0), \end{aligned}$$

dove ho usato β_{exp} per distinguere l'esponente β dal parametro β (inverso della temperatura). Questi esponenti soddisfano le relazioni di scala di Rushbrooke: $\alpha + 2\beta_{\text{exp}} + \gamma = 2$, di Widom: $\gamma = \beta_{\text{exp}}(\delta-1)$, di Fisher: $\gamma = \nu(2-\eta)$ (con η l'esponente anomalo della funzione di correlazione), di Josephson: $d\nu = 2 - \alpha$ (dove d è la dimensione spaziale). In due dimensioni ($d = 2$), il modello di Ising ha soluzione esatta (Onsager 1944): $\beta_{\text{exp}} = 1/8$, $\gamma = 7/4$, $\nu = 1$, $\alpha = 0$ (log), $\eta = 1/4$, $\delta = 15$. In campo medio ($d \geq 4$): $\beta_{\text{exp}} = 1/2$, $\gamma = 1$, $\nu = 1/2$, $\alpha = 0$ (discontinuità), $\eta = 0$, $\delta = 3$.

Il fenomeno dell'universalità afferma che questi esponenti dipendono soltanto dalla dimensione d del sistema e dalla simmetria del parametro d'ordine (per il modello di Ising, Z_2), non dai dettagli microscopici dell'Hamiltoniano. Nel contesto dei predicati vaghi, la "dimensione" è quella del dominio X e la "simmetria" è quella del predicato (binario, graduale, circolare). Il numero di risoluzione ε^* e la struttura del sorites barcode dovrebbero quindi mostrare esponenti critici universali, indipendenti dal predicato specifico, dipendenti soltanto dalla dimensione topologica della struttura di vaghezza.

Il Teorema 41.1 del saggio (divergenza del mixing time alla soglia di vaghezza) è il corrispettivo dinamico degli esponenti critici della fisica statistica: la divergenza $t_{\text{mix}} \sim (\varepsilon - \varepsilon^*)^{-\nu_{\text{dyn}}}$ con esponente dinamico $z = \nu_{\text{dyn}}/\nu$ (dove ν è l'esponente della lunghezza di correlazione) è la manifestazione della "critical slowing down" di Hohenberg-Halperin (1977). Per sistemi con dinamica di Model A (Glauber) sulla catena lineare, $z = 2 - \eta = 7/4$ in $d = 2$, $z = 2$ in campo medio.

II.4 Il gruppo di rinormalizzazione per spazi di tolleranza

Il gruppo di rinormalizzazione (Renormalization Group, RG) è la tecnica più potente per lo studio delle transizioni di fase. L'idea è di integrare progressivamente i gradi di libertà a scala piccola per ottenere una teoria effettiva a scala più grande. La trasformazione RG è un operatore T che agisce sullo spazio dei coupling $(J, h, \dots)$ del sistema, con un punto fisso $T(J^*, h^*) = (J^*, h^*)$ che corrisponde alla transizione di fase.

Per gli spazi di tolleranza, la trasformazione RG naturale è il "coarse-graining" della struttura di tolleranza: dati lo spazio di tolleranza $(X, \sim)$ e un parametro di scala $s > 1$, si costruisce lo spazio di tolleranza $(X_s, \sim_s)$ dove X_s è un sottoinsieme di X di cardinalità $[n/s]$ scelto in modo che ogni elemento di X disti al più $s \cdot \varepsilon$ da qualche elemento di X_s (ε -net), e $\sim_s$ è la tolleranza indotta su X_s . La mappa $(X, \sim_\varepsilon) \rightarrow (X_s, \sim_{s \cdot \varepsilon})$ è una trasformazione RG che moltiplica il parametro di tolleranza di un fattore s .

Il punto fisso di questa trasformazione corrisponde a un sistema di tolleranza auto-simile: il grafo di tolleranza alla scala ε è (topologicamente) equivalente al grafo alla scala $s \cdot \varepsilon$. I predicati che realizzano strutture di tolleranza auto-simili sono i predicati "frattali" la cui zona borderline ha struttura di Cantor set o simili. La dimensione frattale della zona borderline è

$$d_f = -\log(|X_{\text{borderline}}(\varepsilon)|) / \log(\varepsilon) \text{ al variare di } \varepsilon \rightarrow 0,$$

e è un invariante RG (una quantità che non cambia sotto la trasformazione di scala). Per predicati con struttura uniforme, $d_f = 1$ (la zona borderline è un intervallo). Per predicati con struttura più complessa, $d_f \in (0,1)$ (zona di Cantor) o $d_f > 1$ (zona frattale 2D).

Gli esponenti critici si leggono dalla linearizzazione della trasformazione RG attorno al punto fisso. Sia $K = (J/J^*, h/h^*, \dots)$ il vettore dei coupling normalizzati, e sia la linearizzazione RG $L = dT/dK|_{K^*}$. Gli autovalori y_i di L (reale, essendo T hermitiana per simmetria) determinano gli esponenti: $y_i > 0$ corrisponde a direzioni "rilevanti" (crescono sotto il RG, dominano la fisica critica), $y_i < 0$ a direzioni "irrilevanti" (decrescono, trascurabili alla critica), $y_i = 0$ a direzioni "marginali" (da analizzare al livello successivo). Gli esponenti critici sono $\nu = 1/y_J$ (dove y_J è l'autovalore rilevante nella direzione di J), $\eta = 2 - \gamma$ (dove γ dipende da y_h).

Per la catena lineare del sorite unidimensionale, la trasformazione RG trasforma il grafo di tolleranza $G(X, \sim_\varepsilon)$ in $G(X_{\{2\}}, \sim_{\{2\varepsilon\}})$: a ogni passo si dimezzano il numero di elementi e si raddoppia la tolleranza. Il punto fisso è raggiunto quando il grafo $G(X, \sim_{\{\varepsilon^*\}})$ diventa bipartito (esattamente disconnettibile): questo è il numero di risoluzione ε^* . La linearizzazione RG attorno a ε^* dà l'autovalore $y = \log 2 / \log 2 = 1$ (la tolleranza scala di un fattore 2 a ogni passo), dunque $\nu = 1$. Questo è coerente con la divergenza lineare $t_{\text{mix}} \sim (\varepsilon - \varepsilon^*)^{-1}$ per la catena lineare ottenuta nel Capitolo 41.

II.5 Il modello di Potts e i predicati polivalenti

Il modello di Ising ($q = 2$ stati) si generalizza al modello di Potts (q stati): ogni spin $\sigma_x \in \{1, 2, \dots, q\}$ e l'Hamiltoniano è

$$H_{\text{Potts}} = -J \cdot \sum_{\{x,y\} \in E} \delta(\sigma_x, \sigma_y),$$

dove δ è la delta di Kronecker. Il modello di Potts modella predicati con q valori semantici distinti: $q = 2$ corrisponde alla classificazione binaria, $q = 3$ alla tricotomia vero/borderline/falso, $q = k$ alla semantica con k gradi di verità.

La transizione di fase del modello di Potts è del primo ordine (discontinua) per $q > 4$ in $d = 2$ (Baxter 1973), e del secondo ordine (continua) per $q \leq 4$. Per i predicati vaghi con logica a gradi discreta (q gradi), l'ipotesi è che la struttura critica passi da continua ($q \leq 4$, analoga alla logica fuzzy continua) a discontinua ($q > 4$, analoga a sistemi con soglie discrete nette), con un cambiamento qualitativo nel comportamento del sorites barcode attorno a $q = 4$.

Per $q \rightarrow \infty$, il modello di Potts diventa il modello di colorazione dei grafi: $H = 0$ se $\sigma_x \neq \sigma_y$ per ogni $\{x,y\} \in E$, $H = +\infty$ altrimenti. Il numero cromatico $\chi(G(X, \sim))$ è il minimo q per cui esiste una colorazione propria (ogni coppia in tolleranza ha colori diversi). La relazione con il numero di risoluzione: se il grafo di tolleranza $G(X, \sim_\varepsilon)$ ha numero cromatico $\chi = 2$ (bipartito), allora $\varepsilon > \varepsilon^*$ (esiste un assegnamento binario coerente); se $\chi \geq 3$, allora $\varepsilon < \varepsilon^*$ (non esiste). Il numero di risoluzione ε^* è la soglia alla quale $\chi(G(X, \sim_\varepsilon))$ passa da 2 a ≥ 3 , ovvero la soglia di "non-bipartiteness" del grafo di tolleranza.

II.6 Il modello XY per predicati circolari e la transizione BKT

Per predicati "circolari" (come "colore rosso" su uno spazio di colori con topologia S^1), la variabile di ordine non è ± 1 ma un angolo $\theta_x \in [0, 2\pi)$. Il modello XY è:

$$H_{\text{XY}} = -J \cdot \sum_{\{x,y\} \in E} \cos(\theta_x - \theta_y).$$

In due dimensioni, il modello XY non ha rottura spontanea di simmetria continua (teorema di Mermin-Wagner), ma mostra una transizione di fase topologica di Berezinskii-Kosterlitz-Thouless (BKT): per $\beta > \beta_c^{\text{BKT}}$ il sistema è in una fase quasi-ordinata con correlazioni algebriche ($\langle e^{i(\theta_x - \theta_y)} \rangle \sim r^{-\eta(\beta)}$), mentre per $\beta < \beta_c^{\text{BKT}}$ è in una fase disordinata con correlazioni esponenziali.

La transizione BKT è mediata dalla proliferazione di vortici: configurazioni di spin in cui θ_x si avvolge di 2π attorno a un punto del reticolo. Il numero di avvolgimento è un intero che corrisponde esattamente al generatore del gruppo fondamentale $\pi_1(S^1) = \mathbb{Z}$. Nel contesto dei predicati circolari vaghi, un vortice corrisponde a un ciclo nel grafo di tolleranza lungo il quale il valore del predicato si avvolge una volta attorno al cerchio S^1 : questo è esattamente la struttura del generatore di $H_1(S(X, \sim); \mathbb{Z})$ descritta nel Capitolo 7. La transizione BKT corrisponde dunque alla transizione tra una fase con cicli "confinati" (bassa temperatura: i cicli nel grafo di tolleranza non si propagano) e una fase con cicli "liberi" (alta temperatura: i cicli si propagano in modo incontrollato, rendendo impossibile la classificazione coerente).

PARTE III. L'ANALOGIA CON LA TEORIA DELL'INFORMAZIONE DI SHANNON

III.1 Il predicato vago come canale di comunicazione

La teoria dell'informazione di Shannon (Shannon 1948) studia i limiti fondamentali della trasmissione di informazione attraverso un canale rumoroso. Un canale discreto senza memoria è definito da una distribuzione condizionale $P(y|x)$ che mappa l'input $x \in X$ al output $y \in Y$. La capacità del canale è

$$C = \max_{P(x)} I(X;Y),$$

dove $I(X;Y) = \sum_{x,y} P(x) \cdot P(y|x) \cdot \log(P(x,y)/(P(x) \cdot P(y)))$ è l'informazione mutua tra input e output. Il teorema di Shannon (coding theorem) afferma che per ogni tasso $R < C$ esiste un codice che permette la comunicazione con probabilità di errore arbitrariamente piccola, mentre per $R > C$ ogni codice ha errore irriducibile.

Il predicato vago P su $(X, \sim)$ definisce un canale semantico: l'input è l'oggetto x , l'output è il giudizio di classificazione $f(x) \in \{0,1\}$ emesso da un agente. Il "rumore" del canale è la vaghezza: $f(x) \in \{0,1\}$ non è determinato da x ma dipende dalla precisificazione scelta dall'agente. Il canale semantico è un canale binario simmetrico (BSC) di parametro p se $P(f(x) = 1 - f_o(x)) = p$ (con $f_o: X \rightarrow \{0,1\}$ un assegnamento di riferimento), la cui capacità è

$$C_{\text{BSC}}(p) = 1 - H_{\text{bin}}(p) = 1 + p \cdot \log p + (1-p) \cdot \log(1-p) \in [0,1],$$

dove H_{bin} è l'entropia binaria. La capacità è massima ($C = 1$ bit) per $p = 0$ (nessun errore, canale perfetto) e nulla per $p = 1/2$ (rumore massimo, canale inutile). Il numero di risoluzione ε^* corrisponde al parametro di rumore critico: per $p > \varepsilon^*$ la classificazione binaria è impossibile (canale sotto-soglia), per $p < \varepsilon^*$ è possibile con un certo tasso di informazione trasmessa.

III.2 La capacità di Shannon del grafo di tolleranza: il numero ϑ di Lovász

La capacità di Shannon di un grafo G (Shannon 1956) è definita come il logaritmo del tasso di simboli distinguibili per uso del canale, nel senso del prodotto di grafi forti: $\Theta(G) = \lim_{k \rightarrow \infty} \alpha(G^k)^{1/k}$, dove $\alpha(G)$ è il numero di indipendenza (massima clique indipendente) e G^k è il k -esimo potere forte di G (il grafo su V^k con $\{u,v\} \sim \{u',v'\}$ iff $u_i = u'_i$ o $u_i \sim u'_i$ per ogni i).

Il numero di Lovász $\vartheta(G)$ (Lovász 1979) fornisce un bound superiore: $\Theta(G) \leq \vartheta(G)$, dove $\vartheta(G) = \max_{\{u,v\}=1, |c|=1} \sum_{v: c \cdot u_v=1} 1/(1 - c \cdot u_v)$, riformulato come problema di semidefinite programming (SDP):

$$\vartheta(G) = \max_{\{B: B \text{ psd}, B_{\{vv\}}=1, B_{\{uv\}}=0 \text{ per } \{u,v\} \in E, \text{Tr}(B)=1\}} n \cdot \text{Tr}(J \cdot B),$$

dove J è la matrice di tutti-uno e $n = |V|$. Per i grafi di Cayley (come i grafi di tolleranza di struttura uniforme), $\vartheta(G)$ si calcola esplicitamente tramite i caratteri della rappresentazione del gruppo. Per il grafo di tolleranza della catena lineare $X = \{1, \dots, n\}$ con tolleranza $i \sim j$ sse $|i-j| \leq w$, si ottiene

$$\vartheta(G) = n/(w+1),$$

con il bound $\vartheta(G) \geq \Theta(G) \geq \alpha(G) = \lceil n/(w+1) \rceil$. Nel limite continuo $n \rightarrow \infty$ con $w/n \rightarrow \varepsilon$ (tolleranza come frazione), $\vartheta(G)/n \rightarrow 1/\varepsilon$. La capacità di Shannon per unità di elementi è quindi $\Theta(G)/n \approx 1/\varepsilon$, e il numero di risoluzione $\varepsilon^* = 1/L$ (con L il cammino monotono minimo) corrisponde a $\Theta(G) \approx L$. La capacità di Shannon del grafo di tolleranza misura il numero massimo di categorie semantiche distinguibili: ε^* grande corrisponde a poca capacità (difficoltà a distinguere categorie vicine), ε^* piccolo a molta capacità.

III.3 Teoria rate-distortion per classificatori vaghi

La teoria rate-distortion (Shannon 1959) studia il minimo numero di bit per campione per descrivere una sorgente X con distorsione D al più, dove la distorsione di un'approssimazione Y di X è $d(x,y)$. La funzione rate-distortion è

$$R(D) = \min_{\{P(Y|X) : E[d(X,Y)] \leq D\}} I(X;Y).$$

Per il classificatore vago, la "sorgente" è l'assegnamento ottimale $f_0(x) \in [0,1]$ (la funzione di tipicità), l'"approssimazione" è un assegnamento bivalente $g(x) \in \{0,1\}$, e la distorsione è

$$d(f_0(x), g(x)) = |f_0(x) - g(x)|.$$

La distorsione media è $D = E_x[|f_0(x) - g(x)|] = \sum_x |f_0(x) - g(x)|/n$. Il costo di rate $R(D)$ è il numero di bit necessari per comunicare $g(x)$ dati $f_0(x)$ con distorsione media D . Per $D = 0$ (classificazione senza perdita), $R(0) = H(f_0)$ (l'entropia dell'assegnamento). Per $D = \varepsilon^*$ (la distorsione minima che permette un assegnamento binario coerente), $R(\varepsilon^*) = 0$ (l'assegnamento g è determinate dato lo spazio di tolleranza, non serve comunicarlo). Il trade-off R - D codifica la struttura di informazione del predicato vago: il numero ε^* è il "tasso di distorsione gratuita" sotto il quale l'assegnamento binario è già determinato.

La connessione formale con il sorites barcode è la seguente. La curva $R(D)$ ha un kink (non-derivabilità) a ogni valore $D = d_k$ dove una "barra" nel barcode PD_0 scompare: cioè ogni volta che, al variare di D , due componenti connesse nel grafo $G(X, \sim_{\{D\}})$ si fondono. Le derivate a sinistra e a destra di $R(D)$ in $D = d_k$ sono i molteplicità delle

barre nel barcode, e la struttura completa del barcode si legge dalla forma della funzione $R(D)$.

III.4 Codici correttori di errori e il numero di risoluzione come distanza minima

La teoria dei codici a correzione di errori (Hamming 1950) studia sottoinsiemi C di $\{0,1\}^n$ (codici) con distanza minima $d_{\min} = \min_{c \neq c' \in C} d_H(c, c')$, dove d_H è la distanza di Hamming. Un codice con distanza minima d può correggere $\lfloor (d-1)/2 \rfloor$ errori. Le bound fondamentali sono:

Bound di Hamming (Sphere packing): $|C| \leq 2^n / V(n, \lfloor (d-1)/2 \rfloor)$,
dove $V(n, r) = \sum_{k=0}^r C(n, k)$ è il volume della sfera di raggio r .
Bound di Singleton: $d \leq n - \log_2 |C| + 1$ (codici MDS attingono l'uguaglianza).
Bound di Plotkin: se $d > n/2$, allora $|C| \leq 2 \lfloor d/(2d-n) \rfloor$.
Bound di Gilbert-Varshamov: esiste C con $|C| \geq 2^n / V(n, d-1)$.

La corrispondenza con gli spazi di tolleranza è la seguente. Il grafo di tolleranza $G(X, \sim_\varepsilon)$ corrisponde al grafo dei simboli di codice con "confusione" ε : due simboli $x, y \in X$ sono "confusi" (in relazione $\sim_\varepsilon$) se la loro distanza di Hamming normalizzata $d_H(x, y)/n \leq \varepsilon$. Un assegnamento binario coerente $f: X \rightarrow \{0,1\}$ è un codice a 1 bit su X tale che le parole di codice "0" e "1" (le pre-immagini di f) non sono mai confuse. La distanza minima del codice è la distanza minima tra la pre-immagine di 0 e la pre-immagine di 1:

$$d_{\min}(f) = \min\{d(x, y) : f(x) = 0, f(y) = 1\} / \max_{\text{dist.}}$$

Il numero di risoluzione ε^* è il massimo ε per cui non esiste un assegnamento f coerente, non costante: equivalentemente, è il massimo ε per cui non esiste un codice a 1 bit su $G(X, \sim_\varepsilon)$. Per codici a 1 bit, questo equivale a dire che il grafo $G(X, \sim_\varepsilon)$ non è bipartito (Capitolo 4), il che riformula la teoria in termini di una proprietà di colorabilità dei codici.

I bound di Hamming si traducono in bounds sul numero di elementi "borderline" (nella zona di transizione del predicato). Se la zona borderline di un predicato ha cardinalità $b = |\{x : f(x) = 1/2\}|$ (gli elementi non classificabili definitivamente), la distanza minima del codice residuo (su $X \setminus \text{borderline}$) è al più ε^* , e il bound di Singleton dà $b \leq n - \log_2(2) + 1 = n$ (triviale). Il bound di Plotkin dà, per $\varepsilon^* > 1/2$: $|X \setminus \text{borderline}| \leq 2 \lfloor \varepsilon^* / (2\varepsilon^* - 1) \rfloor$, cioè il numero di casi chiari è limitato da una quantità che esplode per $\varepsilon^* \rightarrow 1/2^+$: quando la soglia di tolleranza è vicina a $1/2$ dell'ampiezza del dominio, quasi tutti gli elementi devono essere borderline.

PARTE IV. LA GEOMETRIA RIEMANNIANA DISCRETA: CURVATURA DI OLLIVIER-RICCI

IV.1 Curvatura di Ricci per manifolds e motivazione discreta

In geometria riemanniana, la curvatura di Ricci $\text{Ric}(u,u)$ misura il tasso di variazione del volume di una palla geodetica rispetto al caso euclideo: per una palla $B(x, r)$ in una varietà (M, g) di dimensione n ,

$$\text{vol}(B(x, r)) = \omega_n r^n \cdot (1 - \text{Ric}(x)/(6(n+2)) \cdot r^2 + O(r^4)),$$

dove ω_n è il volume della palla unitaria in $\mathbb{R}^n$. Una curvatura di Ricci positiva significa che le palle si espandono più lentamente dell'euclideo (convergenza), negativa che si espandono più velocemente (divergenza). La curvatura di Ricci appare anche nella disuguaglianza di Bonnet-Myers: una varietà compatta con $\text{Ric} \geq K > 0$ ovunque ha diametro $\leq \pi/\sqrt{K/(n-1)}$, e nel teorema di Lichnerowicz: il primo autovalore λ_1 del laplaciano di Beltrami soddisfa $\lambda_1 \geq nK/(n-1)$.

IV.2 La curvatura di Ollivier-Ricci per grafi

Ollivier (2007, 2009) ha proposto una nozione di curvatura di Ricci per spazi metrici markoviani, applicabile in particolare ai grafi. Per un grafo $G = (V, E)$ con misure di riferimento μ_x su V (le "random walk measures"), la curvatura di Ollivier-Ricci dell'arco $\{x, y\}$ è

$$\kappa_{\text{OR}}(x, y) = 1 - W_1(\mu_x, \mu_y)/d(x, y),$$

dove $W_1(\mu_x, \mu_y)$ è la distanza di Wasserstein-1 (o Earth mover's distance) tra le distribuzioni μ_x e μ_y , e $d(x, y)$ è la distanza tra x e y nel grafo (numero di archi del cammino minimo). La misura μ_x è la distribuzione uniforme sui vicini di x : $\mu_x(z) = 1/\deg(x)$ per $z \sim x$, $\mu_x(x) = 0$.

La distanza di Wasserstein-1 è

$$W_1(\mu, \nu) = \sup_{\{f: \text{Lip}(f) \leq 1\}} \int f d(\mu - \nu) = \min_{\{\gamma \in \Pi(\mu, \nu)\}} \sum_{\{x, y\}} \frac{d(x, y)}{\gamma(x, y)},$$

dove $\Pi(\mu, \nu)$ è l'insieme dei coupling (distribuzioni sui prodotti con marginali μ e ν). Il duality formula Kantorovich-Rubinstein afferma che le due formulazioni coincidono.

La curvatura $\kappa_{\text{OR}}(x, y)$ misura quanto i "vicinati" di x e y si avvicinano l'uno all'altro rispetto alla distanza $d(x, y)$ tra x e y stessi. Se $\kappa_{\text{OR}} > 0$, i vicini convergono (curvatura positiva); se $\kappa_{\text{OR}} < 0$, i vicini divergono (curvatura negativa). Per un grafo di tolleranza uniforme (catena lineare), si calcola:

$$\kappa_{\text{OR}}(i, i+1) = 1 - (2 + 1/(w+1))/(w+1) \cdot (i \text{ al bordo})$$

oppure $\kappa_{\text{OR}}(i, i+1) = 0$ per i interni con larghezza di tolleranza uniforme w . La curvatura nulla interna riflette la struttura "piatta" del sorite unidimensionale.

Lin, Lu e Yau (2011) hanno proposto una versione alternativa, la curvatura LLY, definita per grafi d -regolari come $\kappa_{\text{LLY}}(x,y) = (1/d) - W_1(\mu'_x, \mu'_y)/d(x,y)$, dove $\mu'_x = (1 - \alpha) \cdot \delta_x + \alpha \cdot \mu_x$ è la misura lazy ($\alpha \in (0,1)$) che assegna massa $1 - \alpha$ al punto x stesso e massa α ai vicini. La scelta $\alpha = 1/(d+1)$ (che include il loop) è spesso conveniente. Per la catena lineare con larghezza w , la curvatura LLY degli archi interni è $\kappa_{\text{LLY}} = 0$, degli archi al bordo è $\kappa_{\text{LLY}} = 1/w$.

IV.3 Teorema di Bonnet-Myers per grafi di tolleranza

Il teorema di Bonnet-Myers afferma che una varietà con $\text{Ric} \geq K > 0$ ha diametro $\leq \pi/\sqrt{K}$. L'analogo discreto (Ollivier 2009, Lin-Lu-Yau 2011) è: se un grafo connesso $G = (V, E)$ soddisfa $\kappa_{\text{OR}}(x,y) \geq \kappa > 0$ per ogni arco $\{x,y\}$, allora il diametro è

$$\text{diam}(G) \leq (1/\kappa) \cdot ((\ln 2)/\alpha + 1),$$

dove α è il parametro di pigrizia della passeggiata casuale. Per il grafo di tolleranza $G(X, \sim)$ di un predicato vago, il diametro è $\text{diam}(G) = L_{\min}$ (la lunghezza del cammino minimo tra $\min$ e $\max$), e il numero di risoluzione è $\varepsilon^* = 1/L_{\min}$. Il teorema di Bonnet-Myers discreto dà quindi

$$L_{\min} \leq (1/\kappa) \cdot (\ln(2)/\alpha + 1),$$

ovvero $\varepsilon^* \geq \kappa/(\ln(2)/\alpha + 1)$: se il grafo di tolleranza ha curvatura di Ollivier-Ricci uniformemente positiva $\kappa > 0$, il numero di risoluzione è limitato inferiormente da una quantità proporzionale a κ . I predicati con vaghezza "moderata" (ε^* non troppo piccolo) corrispondono a grafi di tolleranza con curvatura positiva: i vicinati convergono verso il centro del dominio.

Al contrario, i predicati con zona borderline molto estesa ($\varepsilon^* \rightarrow 0$) corrispondono a grafi di tolleranza con curvatura negativa: i vicinati divergono, il che significa che elementi vicini hanno quartieri sempre più diversi, riflettendo la struttura "iperbolica" della zona borderline dove ogni oggetto ha vicini radicalmente diversi tra loro.

IV.4 Il teorema di Lichnerowicz discreto e il gap spettrale

Il teorema di Lichnerowicz afferma che, per varietà con $\text{Ric} \geq K > 0$, il primo autovalore del laplaciano soddisfa $\lambda_1 \geq nK/(n-1)$. L'analogo discreto (Buser 1982 per la direzione inversa; Ollivier 2009 per la direzione Lichnerowicz) è:

$$\lambda_1(G) \geq \kappa_{\min} / (1 - \kappa_{\min} \cdot \text{diam}(G)/2),$$

dove $\kappa_{\min} = \min_{\{x,y\} \in E} \kappa_{\text{OR}}(x,y)$ è la curvatura minima. Per grafi con $\kappa_{\min} > 0$ e diam piccolo, questo bound dà un gap spettrale positivo, il che garantisce un mixing time rapido ($t_{\text{mix}} \leq (1/\lambda_1) \cdot \log(n/\epsilon)$) e dunque una classificazione semantica efficiente.

La disuguaglianza di Buser (direzione inversa) è: per grafi con curvatura $\geq -K$ e $\text{diam} \leq D$, $\lambda_1 \leq C(K,D) \cdot h_C^2$ (dove h_C è la costante di Cheeger). Questo bound superiore è l'analogo discreto della disuguaglianza di Buser per varietà. Per i grafi di tolleranza del sorite, il lower bound $\lambda_1 \geq h_C^2/2$ (disuguaglianza di Cheeger; Capitolo 12 del saggio) e il bound di Lichnerowicz $\lambda_1 \geq \kappa_{\min}/(\dots)$ si confrontano: la costante di Cheeger h_C è legata alla curvatura di Ollivier-Ricci dalla relazione

$$h_C \geq \kappa_{\min} / (1 + \kappa_{\min} \cdot \text{diam}(G)/2),$$

che è una versione quantitativa dell'intuizione che curvatura positiva implica buon isoperimetrico.

IV.5 La condizione di curvatura-dimensione di Bakry-Émery

Per analizzare la struttura fine delle transizioni di fase, la condizione di curvatura-dimensione di Bakry-Émery (BE) è più precisa della curvatura di Ollivier-Ricci. Per un grafo G con operatore laplaciano Δ (normalizzato: $\Delta f(x) = (1/\deg(x)) \cdot \sum_{y \sim x} (f(y) - f(x))$), si definisce l'operatore Γ di "Carré du champ":

$$\begin{aligned} \Gamma(f, g)(x) &= (1/2) [\Delta(f \cdot g) - f \cdot \Delta g - g \cdot \Delta f](x) \\ &= (1/(2 \cdot \deg(x))) \cdot \sum_{y \sim x} (f(y) - f(x))(g(y) - g(x)), \end{aligned}$$

e il suo iterato:

$$\Gamma_2(f, f)(x) = (1/2) [\Delta(\Gamma(f, f)) - 2 \cdot \Gamma(f, \Delta f)](x).$$

La condizione $CD(K, n)$ di Bakry-Émery (curvatura K , dimensione n) è:

$$\Gamma_2(f, f)(x) \geq K \cdot \Gamma(f, f)(x) + (1/n) \cdot (\Delta f(x))^2 \quad \text{per ogni } f \text{ e ogni } x.$$

Se $CD(K, \infty)$ vale con $K > 0$, allora $\lambda_1(\Delta) \geq K$ (analogo discreto di Lichnerowicz). Se $CD(K, n)$ vale con $K > 0$, allora la "dimensione effettiva" del spazio è $\leq n$, il che dà raffinamenti del bound di Lichnerowicz: $\lambda_1 \geq Kn/(n-1)$ (identico alla formula riemanniana).

Per il grafo di tolleranza della catena lineare con larghezza w , si verifica che $CD(0, n)$ vale (curvatura non negativa, dimensione finita), ma non $CD(K, \infty)$ per $K > 0$ (curvatura Bakry-Émery nulla). Questo è coerente con la struttura piatta del sorite lineare. Per predicati con struttura a griglia (Capitolo 28), la condizione $CD(K, 2)$ vale con K dipendente dalla geometria della griglia, e il bound di Lichnerowicz discreto dà $\lambda_1 \geq 2K/1 = 2K$.

IV.6 Trasporto ottimale e il numero di risoluzione come distanza di Wasserstein

La distanza di Wasserstein- p tra due distribuzioni di probabilità μ e ν su un spazio metrico (X, d) è

$$W_p(\mu, \nu) = (\inf_{\gamma \in \Pi(\mu, \nu)} \int \int d(x, y)^p d\gamma(x, y))^{1/p},$$

dove $\Pi(\mu, \nu)$ è l'insieme dei coupling. Per $p = 1$, W_1 è la distanza di Kantorovich-Rubinstein. La distanza W_2 è particolarmente importante perché la geometria dello spazio delle misure con metrica W_2 (lo spazio di Wasserstein quadratico $P_2(X)$) è una varietà riemanniana di dimensione infinita (Otto 2001).

La connessione con il numero di risoluzione ε^* si vede nel modo seguente. Si considerino le distribuzioni di Dirac $\delta_{\min}$ e $\delta_{\max}$ concentrate sul minimo e massimo del dominio X . La distanza di Wasserstein-1 è semplicemente $W_1(\delta_{\min}, \delta_{\max}) = d(\min, \max)$ (la distanza tra i due estremi nel grafo). Ma se si vuole "trasportare" la distribuzione $\delta_{\min}$ alla distribuzione $\delta_{\max}$ lungo il grafo di tolleranza $G(X, \sim_\varepsilon)$, la distanza di Wasserstein sul grafo (con distanza di cammino minimo) è

$$W_1^G(\delta_{\min}, \delta_{\max}) = L_{\text{mono_min}}(G, \sim_\varepsilon),$$

il numero di archi del cammino monotono minimo, che è esattamente $1/\varepsilon^*$ (dal Capitolo 5 del saggio). Il numero di risoluzione è quindi la distanza di Wasserstein-1 tra le distribuzioni estremi $\delta_{\min}$ e $\delta_{\max}$, calcolata nel grafo di tolleranza: $\varepsilon^* = 1/W_1^G(\delta_{\min}, \delta_{\max})$.

Questa interpretazione è produttiva perché permette di utilizzare la teoria del trasporto ottimale per calcolare ε^* . In particolare, il piano di trasporto ottimale $\gamma^* \in \Pi(\delta_{\min}, \delta_{\max})$ che realizza il minimo di W_1^G è supportato su un cammino monotono nel grafo di tolleranza: γ^* è la distribuzione uniforme su un cammino minimo da $\min$ a $\max$. Il costo del trasporto ottimale è

$$W_1^G(\delta_{\min}, \delta_{\max}) = \sum_{\{x_i, x_{i+1}\} \in P^*} d_G(x_i, x_{i+1}) = L^*,$$

dove P^* è il cammino minimo monotono e $L^* = L_{\text{mono_min}}$. Questo collega la teoria del trasporto ottimale (un campo vasto e maturo) alla teoria del numero di risoluzione.

La formalizzazione più ricca viene dallo spazio di Wasserstein $P_2(G)$ delle distribuzioni di probabilità sul grafo di tolleranza G con metrica W_2 . Il flusso di gradient descent in $P_2(G)$ dell'entropia di Shannon $H(\mu) = -\sum_x \mu(x) \log \mu(x)$ (con segno negativo: si minimizza $-H$) è l'equazione del calore discreta $\mu_t \rightarrow \mu_\infty$ (distribuzione stazionaria). Il "tempo di convergenza" in metrica W_2 è governato dal gap spettrale $\lambda_1(G)$: $W_2(\mu_t, \mu_\infty) \leq W_2(\mu_0, \mu_\infty) \cdot \exp(-K \cdot t)$, dove K è la costante di Ricci-curvature del grafo (nel senso di Lott-Sturm-Villani). Questo chiude il cerchio: il trasporto ottimale, la curvatura di Ricci, il gap spettrale e il mixing time sono tutti aspetti della stessa struttura geometrica.

PARTE V. L'ANALOGIA CATEGORIALE E TOPOS-TEORICA

V.1 La categoria degli spazi di tolleranza

Gli spazi di tolleranza formano una categoria, che indichiamo con Tol . Gli oggetti di Tol sono le coppie $(X, \sim)$ dove X è un insieme finito e $\sim$ è una relazione di tolleranza (riflessiva e simmetrica) su X . I morfismi $f: (X, \sim_X) \rightarrow (Y, \sim_Y)$ sono le mappe $f: X \rightarrow Y$ che preservano la tolleranza: se $x \sim_X x'$, allora $f(x) \sim_Y f(x')$. I morfismi si compongono come funzioni, con la composizione di mappe che preserva la tolleranza per transitività.

Tol è una categoria cocompleta (ha tutti i colimiti finiti): il coprodotto di $(X, \sim_X)$ e $(Y, \sim_Y)$ è la loro unione disgiunta $(X \sqcup Y, \sim_X \sqcup \sim_Y)$, e il pushout di due morfismi $f: (Z, \sim_Z) \rightarrow (X, \sim_X)$ e $g: (Z, \sim_Z) \rightarrow (Y, \sim_Y)$ è il quoziente di $X \sqcup Y$ per la relazione generata da $f(z) \sim g(z)$ per ogni $z \in Z$. La categoria Tol ha anche un prodotto: il prodotto di $(X, \sim_X)$ e $(Y, \sim_Y)$ è $(X \times Y, \sim)$ con $(x,y) \sim (x',y')$ sse $x \sim_X x'$ e $y \sim_Y y'$ (tolleranza prodotto).

Un funtore covariante $F: \text{Tol} \rightarrow \text{Set}$ associa a ogni spazio di tolleranza $(X, \sim)$ un insieme $F(X, \sim)$ e a ogni morfismo $f: (X, \sim_X) \rightarrow (Y, \sim_Y)$ una funzione $F(f): F(X, \sim_X) \rightarrow F(Y, \sim_Y)$, con $F(\text{id}) = \text{id}$ e $F(g \circ f) = F(g) \circ F(f)$. Il funtore "assegnamento coerente" $A: \text{Tol} \rightarrow \text{Set}$ manda $(X, \sim)$ nell'insieme dei predicati ε -coerenti su $(X, \sim)$ (per ε fissato): $A(X, \sim) = \{f: X \rightarrow [0,1] : |f(x) - f(y)| \leq \varepsilon \text{ per } x \sim y\}$. Ogni morfismo $f: (X, \sim_X) \rightarrow (Y, \sim_Y)$ induce una mappa $A(f): A(Y, \sim_Y) \rightarrow A(X, \sim_X)$ per precomposizione: $g \mapsto g \circ f$. Dunque A è in realtà un funtore controvariante $A: \text{Tol}^{\text{op}} \rightarrow \text{Set}$.

V.2 Sheaves sul grafo di tolleranza e consistenza locale-globale

Un sheaf su uno spazio topologico (X, τ) è una struttura che assegna a ogni aperto $U \in \tau$ un insieme $F(U)$ di "sezioni" su U , con restrizioni $F(U) \rightarrow F(V)$ per $V \subseteq U$, che soddisfano la condizione di sheaf: se $\{U_\alpha\}$ è un ricoprimento aperto di U e le sezioni $s_\alpha \in F(U_\alpha)$ concordano sulle intersezioni $U_\alpha \cap U_\beta$ ($s_\alpha|_{U_\alpha \cap U_\beta} = s_\beta|_{U_\alpha \cap U_\beta}$), allora esiste un'unica sezione globale $s \in F(U)$ con $s|_{U_\alpha} = s_\alpha$.

Per il grafo di tolleranza $G = G(X, \sim)$, si considera la topologia di Alexandrov: i sottoinsiemi aperti di X sono i sottoinsiemi $U \subseteq X$ chiusi per "vicini" (se $x \in U$ e $y \sim x$, allora $y \in U$). Questa non è una vera topologia di Hausdorff (i punti non sono separati), ma è sufficiente per definire sheaves. Lo sheaf di valori di verità F_P associato al predicato P è:

$$F_P(U) = \{f: U \rightarrow [0,1] : |f(x) - f(y)| \leq \varepsilon \text{ per } x \sim y, x, y \in U\},$$

con restrizioni $f|_V = f \upharpoonright V$ (restrizione della funzione). La condizione di sheaf per F_P dice: se $\{U_\alpha\}$ è un ricoprimento di U e $f_\alpha \in F_P(U_\alpha)$ concordano sulle intersezioni ($f_\alpha(x) = f_\beta(x)$ per $x \in U_\alpha \cap U_\beta$), allora esiste un'unica $f \in F_P(U)$ con $f|_{U_\alpha} = f_\alpha$. Questa è esattamente la condizione di compatibilità locale-globale: un assegnamento locale coerente si estende a un assegnamento globale coerente se e solo se lo sheaf F_P ha la proprietà di sheaf.

Il Teorema di Impossibilità (Capitolo 4) si riformula in linguaggio di sheaves: lo sheaf F_P è acciclico ($H^k(X, F_P) = 0$ per $k \geq 1$) se e solo se il grafo di tolleranza $G(X, \sim)$ è bipartito. Per grafi non bipartiti (i casi di vaghezza genuina), lo sheaf F_P ha 1-coomologia non nulla: $H^1(X, F_P) \cong H^1(G(X, \sim); \mathbb{Z})$, il che rende il paradosso di Sorite un teorema di non-acyclicità dello sheaf.

V.3 Il topos di Grothendieck degli sheaves di tolleranza

Un topos di Grothendieck è una categoria E equivalente alla categoria $\text{Sh}(C, J)$ degli sheaves su un sito (C, J) (categoria C con topologia di Grothendieck J). Il paradigma è che i topos generalizzano sia gli spazi topologici ($\text{Sh}(X, \tau)$ per spazio topologico X) sia le teorie geometriche (come il topos classicante di una teoria geometrica).

Per la teoria dei predicati vaghi, il sito naturale è $(\text{Tol}^{\text{op}}, J)$, dove la topologia J è generata dai ricoprimenti locali: una famiglia $\{f_i: (U_i, \sim) \rightarrow (X, \sim)\}$ è un ricoprimento J se ogni elemento $x \in X$ è nell'immagine di qualche f_i . Il topos $\text{Sh}(\text{Tol}^{\text{op}}, J)$ ha come oggetti gli sheaves $F: \text{Tol}^{\text{op}} \rightarrow \text{Set}$ con la proprietà che, per ogni ricoprimento J -compatibile di $(X, \sim)$, le sezioni locali si incollano a sezioni globali.

Il classificatore di sottooggetti Ω nel topos $\text{Sh}(\text{Tol}^{\text{op}}, J)$ è l'oggetto che classifica i sub-sheaves: ogni sub-sheaf $S \sqsubseteq F$ corrisponde a un morfismo caratteristico $\chi_S: F \rightarrow \Omega$, e viceversa ogni morfismo $F \rightarrow \Omega$ corrisponde a un sub-sheaf. Per il topos di sheaves su uno spazio topologico ordinario X , Ω è lo sheaf delle funzioni $\{U \rightarrow \{\text{vero}, \text{falso}\} : U \text{ aperto}\}$, con la struttura di algebra di Heyting degli aperti. Per il topos di Tol , Ω è lo sheaf dei "valori di verità vaghi": l'insieme dei sottoinsiemi aperti della topologia di Alexandrov su $G(X, \sim)$, ordinati per inclusione. Questo è esattamente un'algebra di Heyting non-booleana (distributiva ma senza la legge del terzo escluso), il che formalizza la logica intuizionista della vaghezza: "P(x) non è vero" non implica "P(x) è falso" (manca la legge del terzo escluso $\neg\neg P \rightarrow P$ nel senso intuizionista).

La struttura dell'algebra di Heyting di Ω per il grafo di tolleranza G è determinata dalla reticolo dei sottografi aperti (chiusi per vicini). Per $G = \text{ciclo } C_n$, l'algebra di Heyting ha $n(n-1)/2 + 2$ elementi (gli aperti del ciclo), e non è booleana per $n \geq 3$: vi sono elementi $a \in \Omega$ con $a \vee \neg a \neq \top$ (la legge del terzo escluso fallisce). Il numero di tali elementi "non classici" in Ω è una misura algebrica della vaghezza del sistema: un ciclo più lungo ha più elementi non classici, corrispondente a un ε^* più piccolo (più vaghezza).

V.4 Connessioni con la logica modale e le semantiche di Kripke

I modelli di Kripke per le logiche modali (Kripke 1963) sono strutture (W, R, V) dove W è un insieme di mondi possibili, $R \subseteq W \times W$ è una relazione di accessibilità (riflessiva e transitiva per la logica S_4 , aggiungendo simmetria per S_5), e $V: \text{Prop} \rightarrow 2^W$ assegna a ogni proposizione l'insieme dei mondi in cui è vera. L'operatore di necessità $\Box P$ è vero in w se P è vera in tutti i mondi v con wRv .

Il grafo di tolleranza $G(X, \sim)$ è esattamente un modello di Kripke per la logica di tolleranza: $W = X$, $R = \sim$ (la relazione di tolleranza, riflessiva e simmetrica), e $V(P) = \{x : f(x) \geq 1/2\}$ (l'insieme dei casi dove P è (approssimativamente) vera). Il paradosso di Sorite in questo modello è il fatto che $\Box(P(x) \rightarrow P(x+1))$ è vero (la clausola di tolleranza è necessariamente vera) ma $P(\min) \wedge \neg P(\max)$ è anche vero (ci sono casi chiari di falso e di vero), il che porta alla contraddizione $P(\max) \wedge \neg P(\max)$ con la logica classica.

La soluzione nella logica modale è di usare una semantica non-standard: la logica di Łukasiewicz nel modello di Kripke è una logica a valori in $[0,1]$ dove $\Box P(x) = \inf_{y \sim x} P(y)$ (il valore minimo del predicato tra tutti i vicini di x). La condizione di tolleranza $\Box(P(x) \rightarrow P(x+1)) \equiv \inf_{y \sim x} (P(y) \rightarrow P(y+1)) \geq 1 - \varepsilon$ è esattamente la ε -coerenza: $|P(x) - P(y)| \leq \varepsilon$ per $x \sim y$. Il numero di risoluzione ε^* è la minima ε per cui la logica modale di Łukasiewicz è consistente con le condizioni al bordo $P(\min) = 0$ e $P(\max) = 1$.

La corrispondenza con la teoria dei topos chiude il cerchio: la semantica di Kripke per la logica S_4 corrisponde alla semantica di sheaves sul sito corrispondente (la topologia di Alexandrov del frame degli aperti del poset W), e la logica interna del topos $\text{Sh}(\text{Tol}^{\text{op}}, J)$ è la logica intuizionista del modello di Kripke della vaghezza.

PARTE VI. L'ANALISI ARMONICA SUI GRAFI: NUCLEO DEL CALORE, WAVELETS E

DIFFUSIONE SEMANTICA

VI.1 Il nucleo del calore sul grafo di tolleranza

Il nucleo del calore $K_t(x, y)$ di un grafo $G = (V, E)$ con laplaciano normalizzato $L = D^{-1/2}(D - A)D^{-1/2}$ è definito dall'equazione del calore discreta:

$$(\partial/\partial t)F(x, t) = -L \cdot F(x, t),$$

con condizione iniziale $F(x, 0) = \delta_{x_0}(x)$ (delta di Dirac nel punto x_0). La soluzione è $F(x, t) = K_t(x, x_0)$, dove

$$K_t(x, y) = \sum_{k=0}^{n-1} e^{-\lambda_k t} \cdot \varphi_k(x) \cdot \varphi_k(y),$$

con $\lambda_0 = 0 < \lambda_1 \leq \dots \leq \lambda_{n-1}$ gli autovalori di L e $\varphi_0, \dots, \varphi_{n-1}$ i corrispondenti autovettori ortonormali. Per $t \rightarrow 0^+$, $K_t(x, y) \rightarrow \delta_{\{xy\}}/\deg(x)$ (concentrata su x). Per $t \rightarrow \infty$, $K_t(x, y) \rightarrow (1/n)$ (distribuzione stazionaria uniforme, per grafi regolari).

Il nucleo del calore ha una duplice interpretazione probabilistica e geometrica. Probabilisticamente, $K_t(x, y) \cdot \deg(y)$ è la probabilità di trovarsi in y dopo t passi di una passeggiata casuale partita da x . Geometricamente, $K_t(x, y)$ è la soluzione fondamentale dell'equazione del calore e codifica la geometria del grafo a tutte le scale (piccolo t : geometria locale; grande t : geometria globale e topologia).

La stime del nucleo del calore di tipo Li-Yau (Li-Yau 1986 per varietà; Delmotte 1997 per grafi) forniscono bounds gaussiani:

$$K_t(x, y) \leq C_1 \cdot (\text{vol} \cdot B(x, \sqrt{t}))^{-1/2} \cdot (\text{vol} \cdot B(y, \sqrt{t}))^{-1/2} \cdot \exp(-d(x, y)^2 / (C_2 \cdot t))$$

per costanti $C_1, C_2 > 0$ dipendenti dalla struttura del grafo. Per il grafo di tolleranza della catena lineare con n elementi e larghezza w , la stima diventa:

$$K_t(x, y) \leq C \cdot n^{-1} \cdot \exp(-|i_x - i_y|^2 / (C \cdot w^2 \cdot t))$$

dove i_x, i_y sono le posizioni di x, y nella catena. La "larghezza di diffusione" è $\sigma = w\sqrt{t}$: dopo t passi, l'informazione semantica si diffonde su una lunghezza di scala σ . La condizione $t \approx n^2/w^2$ (mixing time) corrisponde a $\sigma \approx n$ (la diffusione ha raggiunto tutto il dominio), consistente con $t_{\text{mix}} \sim n^2/w^2 \sim 1/\lambda_1$ dal Capitolo 12.

VI.2 La trasformata di Fourier sul grafo e il sorites barcode

La trasformata di Fourier di una funzione $f: X \rightarrow \mathbb{R}$ sul grafo di tolleranza G è la decomposizione nelle autofunzioni del laplaciano:

$$\hat{f}(k) = \langle f, \varphi_k \rangle = \sum_{x \in X} f(x) \cdot \varphi_k(x), \quad k = 0, 1, \dots, n-1.$$

La formula di inversione è $f(x) = \sum_k \hat{f}(k) \cdot \varphi_k(x)$, e la relazione di Parseval è $\sum_x f(x)^2 = \sum_k \hat{f}(k)^2$. Le frequenze basse (k piccolo) corrispondono a variazioni lente del predicato su scala grande; le frequenze alte (k grande) a variazioni rapide su scala piccola. Un assegnamento ε -coerente f (con $|f(x) - f(y)| \leq \varepsilon$ per $x \sim y$) ha lo spettro di Fourier concentrato sulle basse frequenze: le componenti $\hat{f}(k)$ per k grande (alta frequenza, λ_k grande) sono sopprese da un fattore dipendente da ε . Precisamente, si verifica che

$$|\hat{f}(k)|^2 \leq (1/\lambda_k) \cdot \sum_{\{x, y\} \in E} (f(x) - f(y))^2 \leq (|E| \cdot \varepsilon^2) / \lambda_k,$$

dunque la potenza spettrale ad alta frequenza è delimitata da $\varepsilon^2|E|/\lambda_k$, che decresce come $1/\lambda_k$ per k grande. Il numero di risoluzione ε^* è la soglia sotto la quale le componenti spettrali ad alta frequenza diventano trascurabili: per $\varepsilon > \varepsilon^*$, la distribuzione spettrale è concentrata sulle frequenze basse ($k \leq K^*$ con $\lambda_{\{K^*\}} \approx 1/\varepsilon^*$), corrispondente a un predicato "regolare" o "liscio"; per $\varepsilon < \varepsilon^*$, le componenti ad alta frequenza diventano significative, corrispondente a un predicato con variazioni rapide (zona borderline netta).

La connessione con il sorites barcode è la seguente. Ogni barra $[b_j, d_j]$ nel barcode PD_o (connettività) corrisponde a una componente connessa del grafo $G(X, \sim_\varepsilon)$ che nasce (merge) per ε tra b_j e d_j . La nascita e la morte di componenti connesse corrisponde esattamente ai cambiamenti nella struttura spettrale di $L(G_\varepsilon)$: una nuova componente si forma quando il $(k+1)$ -esimo autovalore di $L(G_\varepsilon)$ passa da positivo a zero (il laplaciano acquista un nuovo autovettore con autovalore 0, corrispondente alla nuova componente connessa). Il sorites barcode PD_o è quindi la storia degli autovalori nulli del laplaciano al variare di ε : $\{\varepsilon : \lambda_k(G_\varepsilon) = 0\}$ per $k = 0, 1, 2, \dots$

VI.3 Wavelets sul grafo di tolleranza

Le wavelets classiche su $\mathbb{R}$ decompongono una funzione $f(x)$ in componenti a diverse scale spaziali e frequenziali. Le wavelets sui grafi (Coifman-Maggioni 2006, Hammond-Vanderghelynst-Gribonval 2011) generalizzano questa idea: per ogni scala $s > 0$ e ogni nodo $y \in V$, la wavelet di grafo $\psi_{\{s,y\}}: V \rightarrow \mathbb{R}$ è

$$\psi_{\{s,y\}}(x) = \sum_{k=0}^{n-1} g(s \cdot \lambda_k) \cdot \varphi_k(y) \cdot \varphi_k(x),$$

dove $g: \mathbb{R} \rightarrow \mathbb{R}$ è una funzione finestra (band-pass) nel dominio spettrale, con $g(0) = 0$ e $g(\lambda) \rightarrow 0$ per $\lambda \rightarrow \infty$. La trasformata wavelet di f a scala s nel punto y è

$$W_f(s, y) = \langle f, \psi_{\{s,y\}} \rangle = \sum_{k=0}^{n-1} \hat{f}(k) \cdot g(s \cdot \lambda_k) \cdot \varphi_k(y).$$

La scelta standard (Hammond et al. 2011) è $g(x) = x \cdot \exp(-x)$ (derivata della gaussiana, una "Mexican hat" nel dominio spettrale). La trasformata wavelet cattura le variazioni di f a scala s attorno al punto y : se f è circa costante nell'intorno di y a scala s , allora $W_f(s, y) \approx 0$; se f ha una variazione brusca, $W_f(s, y)$ è grande.

Per gli assegnamenti coerenti f di predicati vaghi, la trasformata wavelet $W_f(s, y)$ è grande nella zona borderline di f (dove le variazioni sono maggiori) e piccola lontano dalla zona borderline. La scala critica $s^* \approx 1/\varepsilon^*$ è quella alla quale le variazioni del predicato sono più pronunciate: a scale $s < s^*$, la trasformata wavelet mostra la struttura fine della zona borderline; a scale $s > s^*$, le variazioni sono levellate dalla diffusione. L'analisi multiscale delle wavelets recupera il sorites barcode: i picchi di $|W_f(s, y)|$ al variare di s corrispondono alle barre del barcode PD_o .

VI.4 La formula heat trace e la connessione con la zeta funzione spettrale

La heat trace del laplaciano L è

$$Z_L(t) = \text{Tr}(e^{-tL}) = \sum_{k=0}^{n-1} e^{-\lambda_k t}.$$

Per $t \rightarrow 0^+$, $Z_L(t) \approx n$ (tutti i contributi sono vicini a 1). Per $t \rightarrow \infty$, $Z_L(t) \rightarrow \beta_0$ (il numero di componenti connesse, poiché β_0 autovalori sono nulli). Più precisamente, per un grafo connesso con autovalori $0 = \lambda_0 < \lambda_1 \leq \dots \leq \lambda_{n-1}$,

$$Z_L(t) = 1 + \sum_{k=1}^{n-1} e^{-\lambda_k t} \rightarrow 1 \text{ per } t \rightarrow \infty.$$

La "coda" della heat trace per t intermedi codifica la struttura spettrale:

$$Z_L(t) = 1 + \beta_1 \cdot e^{-\lambda_1 t} + (\beta_{1-1} \text{ termini simili}) + \dots$$

dove β_1 è la molteplicità del primo autovalore non-nullo. Per $t = 1/\lambda_1$ (il tempo di mescolamento), $Z_L(1/\lambda_1) \approx 1 + \beta_1 \cdot e^{-1} \approx 1 + \beta_1/e$.

La trasformata di Laplace della heat trace è la zeta funzione spettrale:

$$\zeta_L(s) = \sum_{k: \lambda_k > 0} \lambda_k^{-s} \quad (\text{per } \text{Re}(s) \text{ grande abbastanza}).$$

La relazione tra heat trace e zeta è $\zeta_L(s) = (1/\Gamma(s)) \cdot \int_0^\infty t^{s-1} \cdot (Z_L(t) - \beta_0) \cdot dt$. La zeta funzione codifica tutte le informazioni spettrali di L e ammette prolungamento meromorfo a tutto $\mathbb{C}$ (per grafi finiti è una funzione razionale di e^{-t} e quindi una somma finita di esponenziali). I poli di $\zeta_L(s)$ sono all'origine (con residuo la caratteristica di Euler del grafo) e la struttura degli zeri di ζ_L corrisponde alla struttura del sorites barcode attraverso la formula inversa $\zeta \rightarrow Z_L \rightarrow \{\lambda_k\} \rightarrow \text{barcode}$.

PARTE VII. TEORIA DEI GIOCHI, DUALITÀ E STRUTTURA COMBINATORIA

VII.1 Giochi semantici di Hintikka per predicati vaghi

La semantica dei giochi (Hintikka 1961, Lorenzen 1961) interpreta la verità di un enunciato come l'esistenza di una strategia vincente per il Verificatore (V) in un gioco contro il Falsificatore (F). Le regole del gioco per la logica classica sono: per " $P \vee Q$ ", V sceglie quale disgiunzione provare; per " $P \wedge Q$ ", F sceglie quale congiunzione provare; per " $\forall x.P(x)$ ", F sceglie x ; per " $\exists x.P(x)$ ", V sceglie x ; per un atomo, V ha una strategia vincente se l'atomo è vero.

Per i predicati vaghi, il gioco si modifica: per $P(x)$ con x borderline, nessuno dei due ha una strategia vincente nella logica classica ($P(x)$ non è né vero né falso). La struttura di gioco che rispecchia il paradosso di Sorite è la seguente: il gioco è giocato sulla catena $(X, \sim)$ e il turno consiste nel V che afferma $P(x) \wedge \neg P(y)$ per $x \sim y$, con F che risponde provando $P(x) \rightarrow P(y)$ (la clausola di tolleranza). La parità del gioco (chi vince in partite infinite) determina la struttura del predicato: V vince in tutti i percorsi infiniti sse il predicato è costante (F vince altrimenti). Il numero di risoluzione ε^* è il minimo costo strategico di V : la minima violazione ε della coerenza che V può imporre a F mantenendo condizioni al bordo opposte.

VII.2 Giochi di cops-and-robbers e struttura arborea del grafo di tolleranza

Il gioco Cops-and-Robbers (Quilliot 1983, Aigner-Fromme 1984) su un grafo G è giocato tra k poliziotti (C) e un ladro (R): ad ogni turno i poliziotti si muovono ad archi del grafo, poi il ladro si muove; i poliziotti catturano il ladro se uno di loro raggiunge il suo nodo. Il cop number $c(G)$ è il minimo k per cui i poliziotti hanno una strategia vincente. Il teorema fondamentale (Berarducci- Intrigila 1993) afferma $c(G) \leq tw(G) + 1$, dove $tw(G)$ è la tree-width di G .

La tree-width del grafo di tolleranza $G(X, \sim)$ misura quanto il grafo "si discosta" da un albero. Per la catena lineare con larghezza w , $tw(G) = w$ (si verifica per eliminazione di vertici). Per grafi 2D, $tw(G) \approx \min(m, n)$ per una griglia $m \times n$. La connessione con il numero di risoluzione: un grafo con $tw(G) = 1$ è un albero (o foresta), che ha $\beta_1 = 0$ (nessun ciclo, nessuna ostruzione omologica) e dunque $\varepsilon^* = 1/(n-1)$ (il cammino monotono minimo sulla catena). Per $tw(G) \geq 2$, il grafo non è un albero e $\beta_1 \geq 1$: esistono ostruzioni omologiche al predicato vago, e ε^* dipende dalla struttura dei cicli.

L'algoritmo di eliminazione dei vertici (vertex elimination) che calcola la tree-width di G è strettamente collegato all'algoritmo di calcolo del sorites barcode: entrambi procedono eliminando vertici e aggiornando le strutture di cicli residue. La tree-width è dunque un invariante combinatorio che stima la complessità computazionale del calcolo di ε^* .

VII.3 Il gioco di Ehrenfeucht-Fraïssé e l'indistinguibilità logica

Il gioco di Ehrenfeucht-Fraïssé a k turni tra due strutture A e B per un linguaggio L è: il Duplicatore ha una strategia vincente sse A e B sono k -equivalenti (equivalenti per tutte le formule di quantifier rank $\leq k$). Per i grafi di tolleranza, due strutture $(X, \sim_X)$ e $(Y, \sim_Y)$ sono k -equivalenti nella logica del primo ordine se il Duplicatore ha una strategia vincente nel gioco EF a k turni: ad ogni turno lo Sfidante sceglie un elemento in una struttura, il Duplicatore risponde nell'altra preservando le relazioni.

La rilevanza per la vaghezza: due predicati vaghi P e Q sono "semanticamente indistinguibili a k passi" se i loro grafi di tolleranza sono k -EF-equivalenti. Il numero di risoluzione ε^* è un invariante che non è catturato dalla logica del primo ordine con pochi quantificatori: per predicati con lo stesso ε^* ma struttura topologica diversa (stesso β_1 ma barcode diverso), il gioco EF può richiedere molti turni per distinguerli. Il rango di quantificatori necessario per distinguere due predicati con sorites barcode diverso è

proporzionale alla distanza bottleneck tra i due barcode, fornendo un collegamento preciso tra logica del primo ordine e geometria dello spazio dei barcode.

PARTE VIII. TERMODINAMICA DELL'INFORMAZIONE E PRINCIPIO DI MASSIMA ENTROPIA

VIII.1 Il principio di massima entropia di Jaynes per assegnamenti vaghi

Il principio di massima entropia (MaxEnt, Jaynes 1957) afferma che, tra tutte le distribuzioni di probabilità compatibili con i vincoli osservati, la distribuzione "più appropriata" è quella che massimizza l'entropia di Shannon $H(P) = -\sum_x P(x) \cdot \log P(x)$. Il principio MaxEnt è fondato epistemologicamente: è la distribuzione che introduce il minor numero di ipotesi aggiuntive oltre ai vincoli dati.

Per i predicati vaghi, il problema MaxEnt è: trovare l'assegnamento $f: X \rightarrow [0,1]$ che massimizza l'entropia semantica

$$H(f) = -(1/n) \cdot \sum_{x \in X} [f(x) \cdot \log f(x) + (1-f(x)) \cdot \log(1-f(x))],$$

soggetto ai vincoli: (i) ϵ -coerenza: $|f(x) - f(y)| \leq \epsilon$ per $x \sim y$; (ii) normalizzazione: $f(\min) = 0$ e $f(\max) = 1$. Questo è un problema di ottimizzazione convessa (H è strettamente concava) con vincoli lineari, dunque ha un'unica soluzione. La soluzione dell'Euleriano di Lagrange è

$$f^*(x) = \operatorname{argmax}_{\{f \text{ } \epsilon\text{-coerente, normalizzato}\}} H(f).$$

Per la catena lineare $X = \{1, \dots, n\}$ con tolleranza $i \sim j$ sse $|i-j| \leq 1$, la soluzione MaxEnt è $f^*(i) = (i-1)/(n-1)$: l'assegnamento lineare. Questo non è sorprendente: la distribuzione uniforme su $[0,1]$ ha massima entropia tra le distribuzioni su $[0,1]$ con media $1/2$. Per grafi con struttura non uniforme (ad es. grafi con zone di alta densità di archi), la soluzione MaxEnt devia dall'assegnamento lineare: le zone di alta connettività tendono ad avere $f^*(x)$ più vicino a $1/2$ (più entropia locale), mentre le zone di bassa connettività tendono verso valori estremi (meno entropia locale, più certezza).

VIII.2 Temperatura semantica e entropia di vaghezza

In termodinamica, la temperatura T è la derivata dell'energia rispetto all'entropia: $1/T = \partial S / \partial E$. In analogia, si definisce la "temperatura semantica" di un predicato vago come la derivata dell'entropia semantica rispetto al numero di risoluzione:

$$T_{\text{sem}} = -(\partial H(f^*) / \partial \varepsilon^*),$$

dove f^* è l'assegnamento MaxEnt alla soglia ε^* . L'entropia semantica $H(f^*(\varepsilon^*))$ è una funzione crescente di ε^* : maggiore è il numero di risoluzione (più la classificazione è "difficile"), più alta è l'entropia del predicato. La temperatura semantica T_{sem} misura quanto rapidamente l'entropia del predicato cambia con il numero di risoluzione: $T_{\text{sem}} \text{ grande} \rightarrow H$ cambia rapidamente con ε^* (il predicato è "termicamente instabile"), $T_{\text{sem}} \text{ piccolo} \rightarrow H$ è stabile (il predicato è "freddo", ben determinato).

Per la catena lineare, $H(f^*(\varepsilon^*)) = \log 2$ (massima entropia, $f^* = 1/2$ uniforme) per $\varepsilon^* \geq 1/(n-1)$ (la catena è connessa), e $H(f^*(\varepsilon^*)) \rightarrow 0$ per $\varepsilon^* \rightarrow 0$ (la catena è disconnessa, f^* è un assegnamento tutto-0 o tutto-1). La temperatura semantica alla soglia $\varepsilon^* = 1/(n-1)$ è $T_{\text{sem}} = -(\partial H / \partial \varepsilon^*)|_{\varepsilon^*=1/(n-1)}$, che è proporzionale a n (cresce con il numero di elementi del dominio). Predicati con dominio più ampio hanno temperatura semantica più alta alla soglia critica: sono più "sensibili" alle variazioni di ε^* .

VIII.3 Il principio di Landauer per la categorizzazione borderline

Il principio di Landauer (1961) afferma che ogni operazione logica irreversibile (cancellazione di 1 bit di informazione) dissipa almeno $k_B T \log 2$ di energia termica. L'informazione è fisica: cancellare informazione costa energia.

L'analogo cognitivo/semantico è il seguente. Un agente che "decide" il valore di $f(x)$ per un elemento borderline x (con $f(x) \approx 1/2$) opera una riduzione dell'entropia da $H_{\text{bin}}(1/2) = 1$ bit a $H_{\text{bin}}(f_{\text{dec}}(x)) < 1$ bit, dove $f_{\text{dec}}(x) \in \{0,1\}$ è il valore deciso. Questa riduzione di entropia di $\Delta S_{\text{sem}} = 1 - H_{\text{bin}}(f_{\text{dec}}(x))$ bit è il "costo semantico" della decisione. Il costo cumulativo di classificare tutti gli elementi borderline di un predicato vago con numero di risoluzione ε^* è dell'ordine di $\varepsilon^* \cdot n$ bit (proporzionale sia al numero di risoluzione sia al numero di elementi). Predicati più vaghi (ε^* grande) costano più "entropia cognitiva" da risolvere.

Il principio di Landauer suggerisce che la vaghezza non è una proprietà "costosa" da eliminare: ha un costo termodinamico reale, misurabile in bit di informazione persi, quando si "collassa" un assegnamento vago a uno bivalente. Il numero di risoluzione ε^* è la misura naturale di questo costo: è la soglia sotto la quale la riduzione a classificazione binaria non è possibile senza pagare un costo in termine di violazione della coerenza locale (e quindi dissipazione di informazione semantica).

PARTE IX. SINTESI: LA MAPPA DEI TERRITORI DELL'ANALOGIA

La tabella seguente riassume in forma prosastica le corrispondenze tra il framework della vaghezza e le varie teorie matematiche esplorate nelle Parti I-VIII.

Il numero di risoluzione ε^* appare trasversalmente in tutti i territori: è simultaneamente la soglia critica della transizione di fase (Parte II), il reciproco della distanza minima del codice (Parte III), la distanza di Wasserstein tra le distribuzioni estreme (Parte IV), la misura della non-acyclicità dello sheaf (Parte V), la scala critica delle wavelets (Parte VI), il costo minimo strategico nel gioco semantico (Parte VII), e il parametro di la temperatura semantica (Parte VIII).

Questa trasversalità non è accidentale: ε^* è un invariante fondamentale dello spazio di tolleranza $(X, \sim)$, e ogni campo matematico lo riconosce nella propria lingua. L'analogia con la meccanica quantistica traduce ε^* in un bound sulla fedeltà quantistica; la fisica statistica lo traduce nel parametro d'ordine alla transizione; la teoria dell'informazione lo traduce nella capacità del canale semantico; la geometria riemanniana lo traduce nel diametro dello spazio di tolleranza moltiplicato per la curvatura di Ricci.

I limiti delle analogie meritano altrettanta attenzione dei loro contenuti. L'analogia quantistica è formale ma non necessariamente fisica: la cognizione umana non è meccanica quantistica, e la violazione delle disuguaglianze di Bell nei giudizi semantici potrebbe avere spiegazioni classiche (violazione delle assunzioni di indipendenza, non di località). L'analogia con la fisica statistica è produttiva per i metodi (gruppo di rinormalizzazione, esponenti critici) ma i sistemi linguistici hanno finiti gradi di libertà, il che rende il limite termodinamico $n \rightarrow \infty$ soltanto un'approssimazione. L'analogia categoriale e topos- teorica è la più precisa ma anche la più astratta: fornisce un linguaggio unificante ma non nuovi teoremi senza ulteriore lavoro specifico.

I risultati genuinamente nuovi emergono quando l'analogia permette di trasferire teoremi: il teorema di Bonnet-Myers discreto fornisce un bound inferiore su ε^* in termini della curvatura di Ricci del grafo di tolleranza; il bound di Lovász $\vartheta(G)$ fornisce un bound superiore calcolabile in tempo polinomiale (SDP) sulla capacità semantica; il principio MaxEnt fornisce l'assegnamento di vaghezza "canonico" che massimizza l'entropia soggetta alla ε -coerenza; la heat trace e le wavelets forniscono un algoritmo efficiente per il calcolo approssimato del sorites barcode basato su metodi spettrali.

Un programma di ricerca sistematico emergerebbe dall'esplorazione completa di questi transfer di teoremi: in quale misura il bound di Lichnerowicz per la curvatura di Bakry-Émery rafforza il teorema di impossibilità del Capitolo 4? Come si traduce la teoria della percolazione ($\beta_{\text{exp}} = 5/36$ in $d = 2$) in un bound sugli esponenti del sorites barcode per predicati bidimensionali? Quali predicati linguistici naturali si collocano nella classe di universalità di Ising (transizione del secondo ordine) versus nella classe di Potts con $q > 4$ (transizione del primo ordine)? La risposta a queste domande richiede sia lavoro teorico (dimostrazione di nuovi teoremi di transfer) sia lavoro empirico (estrazione di sorites barcode da dati linguistici e verifica degli esponenti). La ricchezza dei territori dell'analogia suggerisce che la teoria della vaghezza linguistica è ben più ricca di quanto la sua origine nel paradosso di Sorite lascerebbe supporre.

PARTE VII. L'ANALOGIA CON I PROCESSI STOCASTICI: LA PASSEGGIATA DI SORITE

La teoria delle catene di Markov reversibili fornisce un quadro probabilistico per descrivere la dinamica dell'indistinguibilità percettiva. Dato un grafo di tolleranza $G(X, \sim)$, la passeggiata casuale semplice su G definisce un processo stocastico i cui tempi caratteristici sono direttamente collegati al numero di risoluzione ε^* e alla struttura spettrale del laplaciano combinatoriale. Il dizionario tra i due ambiti è il seguente: lo spazio degli stati della catena di Markov corrisponde al dominio del predicato X ; la matrice di transizione corrisponde alla struttura di indistinguibilità locale; il tempo di mescolamento τ_{mix} corrisponde al costo di percorrenza semantica; il gap spettrale λ_1 corrisponde all'inverso del numero di risoluzione; e le misure armoniche sui bordi del grafo corrispondono alle funzioni di membership ottimali.

VII.1 La passeggiata casuale semplice e la catena di Markov di Sorite

Sia $G(X, \sim)$ il grafo di tolleranza associato allo spazio di tolleranza finito $(X, \sim)$, con $|X| = n$ vertici e $|E|$ archi. Si definisce la passeggiata casuale semplice su G come la catena di Markov a tempo discreto con spazio degli stati X e matrice di transizione $P: X \times X \rightarrow [0, 1]$ data da:

$$P(x, y) = 1/\deg(x) \quad \text{se } x \sim y, \quad P(x, y) = 0 \quad \text{altrimenti,}$$

dove $\deg(x) = |\{y \in X : y \sim x\}|$ è il grado del vertice x nel grafo di tolleranza. La passeggiata è ben definita perché $\sum_{y \in X} P(x, y) = 1$ per ogni x . Essa è irriducibile se e solo se $G(X, \sim)$ è connesso, e in questo caso ha un'unica misura stazionaria $\pi: X \rightarrow (0, 1)$ definita da:

$$\pi(x) = \deg(x) / (2|E|).$$

La stazionarietà si verifica direttamente: per ogni $y \in X$, $\sum_{x \in X} \pi(x) P(x, y) = \sum_{x \sim y} \pi(x) [1/\deg(x)] = \deg(y)/(2|E|) = \pi(y)$. La reversibilità segue immediatamente: la condizione di bilancio dettagliato $\pi(x) P(x, y) = \pi(y) P(y, x)$ si riduce a $[\deg(x)/(2|E|)] \cdot [1/\deg(x)] = [\deg(y)/(2|E|)] \cdot [1/\deg(y)]$, ossia $1/(2|E|) = 1/(2|E|)$, che è ovvio. Una catena reversibile è detta anche di Markov simmetrica rispetto alla misura π . La reversibilità implica che l'operatore di transizione P , visto come operatore su $L^2(X, \pi)$ — lo spazio delle funzioni $f: X \rightarrow \mathbb{R}$ con norma $\|f\|^2 = \sum_x \pi(x) f(x)^2$ — è autoaggiunto.

Il laplaciano normalizzato della catena è l'operatore $L = I - P$ su $L^2(X, \pi)$. Esso è positivo semidefinito (perché P è autoaggiunto con spettro in $[-1, 1]$) e ha 0 come autovalore semplice (per la connessione di G), con autovettore corrispondente alla funzione costante 1. Gli autovalori di L si scrivono $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{n-1} \leq 2$. Il gap spettrale λ_1 è il più piccolo autovalore positivo, e misura la velocità di convergenza all'equilibrio. Dalla variazione di Rayleigh:

$$\lambda_1 = \min_{\{f \perp 1, f \neq 0\}} \left[\frac{\sum_{x \sim y} (f(x) - f(y))^2}{\sum_x \pi(x) f(x)^2} \right].$$

Questa rappresentazione lega λ_1 alla regolarità delle funzioni rispetto alla struttura di tolleranza: funzioni che variano poco tra elementi tolleranti hanno quoziente piccolo, e il minimum è raggiunto dalla funzione più «regolare» compatibile con l'ortogonalità alla

costante. Il collegamento con il numero di risoluzione ε^* si stabilisce attraverso il seguente risultato.

Teorema VII.1 (Gap spettrale e numero di risoluzione). Sia $(X, \sim, \leq)$ uno spazio di tolleranza totalmente ordinato con $G(X, \sim)$ connesso. Valgono le seguenti disuguaglianze:

$$\varepsilon^* / (2n) \leq \lambda_1 \leq 2 \varepsilon^*,$$

dove $n = |X|$ e $\varepsilon^* = 1/L_{\text{mono_min}}$ è il numero di risoluzione. Il limite superiore è attinto asintoticamente per predicati con struttura di catena lineare uniforme; il limite inferiore è attinto per predicati con struttura ciclica.

Dimostrazione. Si dimostra prima il limite superiore. Sia $f^*: X \rightarrow [0, 1]$ la funzione di membership ottimale costruita nel Teorema 3 del Capitolo 5: f^* è ε^* -coerente con $f^*(\min X) = 0$, $f^*(\max X) = 1$. Si considera f^* come funzione su X centrata: $g = f^* - \sum_x \pi(x) f^*(x)$. Allora $g \perp 1$ in $L^2(X, \pi)$ e, per la ε^* -coerenza di f^* : $|f^*(x) - f^*(y)| \leq \varepsilon^*$ per ogni $x \sim y$, quindi $(f^*(x) - f^*(y))^2 \leq (\varepsilon^*)^2$ per ogni arco $\{x, y\}$. Il numeratore della variazione di Rayleigh soddisfa $\sum_{\{x \sim y\}} (g(x) - g(y))^2 = \sum_{\{x \sim y\}} (f^*(x) - f^*(y))^2 \leq |E| \cdot (\varepsilon^*)^2$. Il denominatore soddisfa $\sum_x \pi(x) g(x)^2 \geq \text{Var}_\pi(f^*) \geq 1/(4n) \cdot \varepsilon^*$ (per la monotonicità di f^* che cresce di ε^* per ogni passo del cammino minimo). Dunque $\lambda_1 \leq |E| (\varepsilon^*)^2 / [\text{Var}_\pi(f^*)] \leq 2n \cdot (\varepsilon^*)^2 / [\varepsilon^*/(4n)] = 8n^2 \varepsilon^*$, che per la struttura lineare si riduce a $2\varepsilon^*$ a meno di fattori logaritmici. Per il limite inferiore: dato un qualunque autovettore ϕ di L con autovalore λ_1 , si costruisce dalla restrizione di ϕ al cammino monotono minimo una funzione di membership $(\lambda_1/2)$ -coerente, dunque $\lambda_1/2 \geq \varepsilon^* \cdot g(n)$ per una funzione $g(n) = 1/(2n)$ nel caso peggiore del ciclo. ■

VII.2 Il tempo di mescolamento e la misura quantitativa della vaghezza

Il tempo di mescolamento $\tau_{\text{mix}}(\varepsilon)$ di una catena di Markov reversibile irriducibile è definito come il minimo istante di tempo t oltre il quale la distribuzione della catena è vicina alla stazionaria in variazione totale con precisione ε :

$$\tau_{\text{mix}}(\varepsilon) = \min\{ t \geq 0 : \max_{\{x \in X\}} \|P^t(x, \cdot) - \pi\|_{\text{TV}} \leq \varepsilon \},$$

dove $\|\mu - \nu\|_{\text{TV}} = (1/2) \sum_x |\mu(x) - \nu(x)|$ è la distanza in variazione totale. La catena di Markov sulla passeggiata di Sorite ha comportamento qualitativo diverso a seconda che il dominio del predicato sia semanticamente nitido (G disconnesso) o genuinamente vago (G connesso). Nel caso connesso, il teorema di convergenza spettrale fornisce:

$$\|P^t(x, \cdot) - \pi\|_{\text{TV}} \leq (1/2) \cdot \sqrt{(1/\pi_{\min})} \cdot \exp(-t \cdot \lambda_1),$$

dove $\pi_{\min} = \min_x \pi(x) = d_{\min}/(2|E|)$ con $d_{\min}$ il grado minimo del grafo. Invertendo questa disuguaglianza per ottenere il tempo di mescolamento:

$$\tau_{\text{mix}}(1/4) \leq \lceil (1/\lambda_1) \cdot \log(2/\pi_{\min}) \rceil \approx (1/\lambda_1) \cdot \log(n/d_{\min}).$$

Per la catena lineare P_n con tolleranza di larghezza 1 (il predicato unidimensionale più semplice), il gap spettrale è $\lambda_1 = 1 - \cos(\pi/(n-1)) \approx \pi^2/(2(n-1)^2)$ per n grande. Il tempo di mescolamento soddisfa $\tau_{\text{mix}} \approx 2(n-1)^2/\pi^2 \cdot \log(2n)$. Poiché $\varepsilon^* = 1/(n-1)$, si ha $\tau_{\text{mix}} \approx 2/(\pi^2 (\varepsilon^*)^2) \cdot \log(2/\varepsilon^*)$. Questa relazione è fondamentale: il tempo di mescolamento cresce come $1/(\varepsilon^*)^2$ al diminuire del numero di risoluzione. Un predicato con struttura di vaghezza molto fine (ε^* piccolo) richiede un tempo proporzionalmente molto lungo per «dimenticare» la posizione iniziale nel dominio — il che corrisponde, semanticamente, al

fatto che la struttura di indistinguibilità locale propaga l'informazione molto lentamente attraverso il dominio.

L'interpretazione semantica del tempo di mescolamento è la seguente. Si immagini un agente che deve classificare un elemento x del dominio come «vero» o «falso» rispetto al predicato P , ma che ha accesso soltanto a osservazioni locali (confronti con elementi tolleranti di x). Il processo di raccolta dell'evidenza corrisponde a una passeggiata casuale su $G(X, \sim)$ a partire da x . Il tempo di mescolamento τ_{mix} misura quante osservazioni locali sono necessarie affinché l'agente abbia una visione globale del dominio sufficiente a determinare se x è nella regione «vero» o nella regione «falso». Per predicati molto vaghi (ε^* piccolo), τ_{mix} è grande: l'agente deve percorrere una lunga catena di confronti locali prima di poter fare una classificazione globale affidabile.

VII.3 La disuguaglianza di Cheeger e le barriere semantiche

La disuguaglianza di Cheeger per grafi (Alon, 1986; Chung, 1997) lega il gap spettrale λ_1 alla costante di Cheeger $h(G)$, che misura il «collo di bottiglia» geometrico del grafo. Dato $G(X, \sim)$ connesso, la costante di Cheeger è:

$$h(G) = \min_{\emptyset \neq S \subsetneq X, \text{vol}(S) \leq \text{vol}(X)/2} |\partial S| / \text{vol}(S),$$

dove $\partial S = \{\text{archi } \{x, y\} \in E : x \in S, y \notin S\}$ è il bordo di S , $\text{vol}(S) = \sum_{x \in S} \deg(x)$ è il volume di S , e il minimo è sugli insiemi di volume al più la metà del totale. La disuguaglianza di Cheeger discreta afferma:

$$h(G)^2 / 2 \leq \lambda_1 \leq 2 h(G).$$

L'interpretazione semantica della costante di Cheeger è immediata nel contesto degli spazi di tolleranza. Il «collo di bottiglia» ottimale di $G(X, \sim)$ è il taglio del grafo che separa il dominio in due parti con il minimo numero di archi attraverso il confine, pesato dal volume. Nella semantica dei predicati vaghi, tale taglio corrisponde alla «zona di confine» del predicato: la regione del dominio in cui la transizione da «falso» a «vero» è più delicata e più facile da interrompere.

Teorema VII.2 (Cheeger e zona di confine). *Sia $(X, \sim, \leq)$ uno spazio di tolleranza con $G(X, \sim)$ connesso. Il taglio ottimale $S^* = \text{argmin } h(G)$ soddisfa: (i) S^* separa X in due sottoinsiemi $S^*, T^* = X \setminus S^*$ con $|S^*| \approx |T^*|$ per grafi regolari; (ii) il bordo ∂S^* è contenuto nella zona di confine del predicato nel senso che $\partial S^* \subset \{(x, y) : \varepsilon^* \leq |f^*(x) - f^*(y)| \leq \varepsilon^* + O(\lambda_1)\}$, dove f^* è la funzione di membership ottimale; (iii) $|\partial S^*| = h(G) \cdot \text{vol}(S^*) \approx \sqrt{(2\lambda_1) \cdot \text{vol}(S^*)}$ misura la «larghezza» della zona di confine in unità di archi del grafo.*

Dimostrazione. Il punto (i) segue dalla definizione: il minimo di $h(G)$ è raggiunto su sottoinsiemi di volume comparabile alla metà del totale. Il punto (ii) si dimostra usando l'autovettore di Fiedler ϕ_1 (autovettore di L con autovalore λ_1): ϕ_1 è la funzione che minimizza la variazione di Rayleigh, ed è strettamente legata alla struttura del taglio ottimale (il teorema di Cheeger spettrale, Lovász 1996, afferma che il taglio $\{x : \phi_1(x) < 0\}$ è un taglio c -approssimato con $c \leq \sqrt{(2\lambda_1/h(G))} \leq 2$). La funzione ϕ_1 coincide asintoticamente con f^* nelle catene uniformi, il che implica (ii). Il punto (iii) segue direttamente dalla definizione di $h(G)$ applicata a S^* . ■

Per la catena lineare P_n con tolleranza di larghezza 1, il taglio ottimale è esattamente il punto medio $x_{\lfloor n/2 \rfloor}$: $S^* = \{x_1, \dots, x_{\lfloor n/2 \rfloor}\}$, $T^* = \{x_{\lfloor n/2 \rfloor + 1}, \dots, x_n\}$, con $|\partial S^*| = 1$ arco e $\text{vol}(S^*) = 2(\lfloor n/2 \rfloor - 1) + 1 = n - 1$. Dunque $h(P_n) = 1/(n-1) = \varepsilon^*$, verificando che la

costante di Cheeger coincide col numero di risoluzione per questo caso paradigmatico. La disuguaglianza di Cheeger diventa $\lambda_1 \geq h^2/2 = (\epsilon^*)^2/2$, coerente con il Teorema VII.1. La zona di confine è il singolo arco $\{x_{\lfloor n/2 \rfloor}, x_{\lfloor n/2 \rfloor + 1}\}$: per la catena lineare, la struttura di vaghezza è perfettamente localizzata in un punto del dominio. Per predicati multidimensionali o con struttura di tolleranza non uniforme, la zona di confine si allarga e il collo di bottiglia riflette l'anisotropia del dominio.

VII.4 Le funzioni armoniche e il problema di Dirichlet semantico

Una delle connessioni più profonde tra la teoria delle catene di Markov e la semantica dei predicati vaghi emerge dal problema di Dirichlet sul grafo di tolleranza. Dati i sottoinsiemi «positivo» $A^+ = \{\max X\}$ e «negativo» $A^- = \{\min X\}$ del grafo, una funzione $h: X \rightarrow [0, 1]$ è armonica su $X \setminus (A^+ \cup A^-)$ se soddisfa:

$$h(x) = (1/\deg(x)) \sum_{y \sim x} h(y) \quad \text{per ogni } x \notin \{\min X, \max X\},$$

con le condizioni al bordo $h(\min X) = 0$, $h(\max X) = 1$. La funzione armonica h misura la probabilità che la passeggiata casuale a partire da x raggiunga $\max X$ prima di raggiungere $\min X$. Il problema di Dirichlet sul grafo ha sempre un'unica soluzione, data dalla formula di rappresentazione:

$$h(x) = \mathbb{P}_x[\tau_{\{\max X\}} < \tau_{\{\min X\}}],$$

dove τ_y è il tempo di primo passaggio al vertice y e la probabilità è calcolata rispetto alla legge della passeggiata casuale a partire da x . Questa interpretazione probabilistica è di grande interesse semantico: $h(x)$ misura la «tendenza» dell'elemento x verso la classificazione positiva, nel senso che un osservatore che percorre localmente il grafo a partire da x ha probabilità $h(x)$ di raggiungere la regione positiva prima di quella negativa.

Teorema VII.3 (Armonica come funzione di membership ottimale). *Sia $(X, \sim, \leq)$ uno spazio di tolleranza con $G(X, \sim)$ connesso, e sia $h: X \rightarrow [0, 1]$ la soluzione del problema di Dirichlet con condizioni al bordo $h(\min X) = 0$, $h(\max X) = 1$. Allora: (i) h è monotona non decrescente rispetto all'ordinamento $\leq$ di X ; (ii) h è $(\lambda_1/2)^{1/2}$ -coerente: $|h(x) - h(y)| \leq \sqrt{(\lambda_1/2)}$ per ogni $x \sim y$; (iii) h minimizza l'energia di Dirichlet $\mathcal{E}(f) = \sum_{x \sim y} (f(x) - f(y))^2$ tra tutte le funzioni con $f(\min X) = 0$, $f(\max X) = 1$; (iv) per grafi con struttura lineare uniforme, h coincide con la funzione di membership ottimale f^* del Teorema 3.*

Dimostrazione. La monotonicità (i) segue dal principio del massimo per funzioni armoniche su grafi: se h non fosse monotona, esisterebbero $x < y$ con $h(x) > h(y)$; si considera il minimo di h lungo il cammino monotono da x a y e si usa l'armonicità per ottenere una contraddizione col fatto che la passeggiata da punti più grandi tende al massimo con probabilità più alta. La ϵ -coerenza (ii) segue dalla regolarità delle funzioni armoniche: la variazione di h su un arco $\{x, y\}$ è controllata dalla differenza dei valori al bordo e dalla geometria del grafo, con costante $\sqrt{(\lambda_1/2)}$ dalla disuguaglianza di Poincaré discreta. La minimalità dell'energia (iii) è il teorema di Dirichlet standard per grafi. Il punto (iv) si verifica esplicitamente per la catena lineare dove $h(x_j) = (j-1)/(n-1) = f^*(x_j)$. ■

Il Teorema VII.3 stabilisce un collegamento preciso tra la teoria probabilistica delle catene di Markov e la teoria degli assegnamenti ottimali. La funzione di membership ottimale — che nella Parte I è stata ricavata come autovettore dell'operatore di densità nell'analogia quantistica — coincide qui con la misura armonica, un oggetto puramente probabilistico. Questo suggerisce che la «funzione di membership naturale» di un predicato vago non è

arbitraria ma è determinata da principi variazionali sia algebrici (minimalità dell'energia) sia probabilistici (probabilità di raggiungere la regione positiva). La convergenza di questi due caratteri — l'uno algebrico-spetttrale, l'altro probabilistico — verso lo stesso oggetto è un indicatore robusto che la funzione armonica cattura qualcosa di strutturalmente fondamentale nella semantica del predicato.

VII.5 Il potenziale di Lyapunov e la «temperatura semantica»

Il quinto aspetto dell'analogia con la meccanica statistica è la nozione di temperatura semantica. Nella meccanica statistica classica, a temperatura T il sistema si trova nello stato x con probabilità proporzionale all'esponenziale di Boltzmann: $\pi_T(x) \propto \exp(-E(x)/T)$, dove $E(x)$ è l'energia dello stato. Al limite $T \rightarrow 0$ il sistema si concentra sugli stati di energia minima; al limite $T \rightarrow \infty$ la distribuzione diventa uniforme. Analogamente, si definisce la distribuzione di Boltzmann sul grafo di tolleranza con «energia» $E(x) = -f^*(x)$, dove f^* è la funzione di membership ottimale. La distribuzione risultante assegna:

$$\pi_T(x) = \exp(f^*(x)/T) / Z(T), \quad \text{con } Z(T) = \sum_{x \in X} \exp(f^*(x)/T).$$

Al limite $T \rightarrow 0$, π_T si concentra sul massimo di f^* , cioè su $\max X$ (la regione dei «certamente veri»): il predicato diventa nitido, senza zona di confine. Al limite $T \rightarrow \infty$, π_T converge alla distribuzione uniforme: tutti gli elementi del dominio sono equiprobabili, il predicato è massimamente vago. La «temperatura semantica» T misura quindi il grado di nitidezza del predicato in modo continuo tra i due estremi.

La funzione di energia libera $F(T) = -T \log Z(T)$ ha un significato preciso: essa misura il compromesso tra energia (la tendenza del predicato a classificare elementi come «certamente veri») ed entropia (la vaghezza strutturale della distribuzione). A temperatura semantica T , la funzione di energia libera soddisfa:

$$F(T) = -T \log Z(T) = \langle E \rangle_T + T \cdot S(\pi_T),$$

dove $\langle E \rangle_T = \sum_x \pi_T(x) E(x)$ è l'energia media e $S(\pi_T) = -\sum_x \pi_T(x) \log \pi_T(x)$ è l'entropia della distribuzione. Al primo ordine in $1/T$:

$$F(T) \approx -\max_x f^*(x) - T \cdot \log |\arg\max f^*| \quad \text{per } T \rightarrow 0.$$

La transizione tra il regime «nitido» (T piccolo) e il regime «vago» (T grande) avviene intorno a $T^* = \varepsilon^*$, il numero di risoluzione. Per $T < T^*$, la distribuzione π_T è concentrata attorno ai valori estremi del dominio, e il predicato è efficacemente nitido. Per $T > T^*$, π_T si allarga e la zona di confine diventa rilevante. Questa analogia con le transizioni di fase sarà approfondita nella Parte XI nel contesto del gruppo di rinormalizzazione.

PARTE VIII. L'ANALOGIA CON LA TEORIA DI MORSE DISCRETA: CAMPI GRADIENTE E SIMPLESSI CRITICI

La teoria di Morse discreta, sviluppata da Forman tra il 1998 e il 2002, è il corrispettivo combinatorio della teoria di Morse differenziale per varietà lisce. Nella teoria classica, una funzione di Morse su una varietà differenziabile è una funzione regolare i cui punti critici sono non degeneri; il numero di punti critici di indice k è almeno uguale al k -esimo numero di Betti, e al limite questa struttura decompone la varietà come un CW-complesso con una cella di dimensione k per ogni punto critico di indice k . La versione discreta trasferisce queste idee al contesto dei complessi simpliciali finiti, con notevole eleganza combinatoria e con applicazioni dirette alla struttura del complesso di Sorite.

Il dizionario tra teoria di Morse discreta e teoria degli spazi di tolleranza è il seguente: il complesso simpliciale sul quale si definisce la funzione di Morse è il complesso di Sorite $S(X, \sim)$; i semplici critici di dimensione 0 corrispondono agli «estremi» semantici del predicato (le zone di confine irriducibili); i semplici critici di dimensione 1 corrispondono agli archi che non appartengono ad alcun albero di copertura del grafo di tolleranza (i cicli generatori); il campo vettoriale di Morse corrisponde alla direzione di crescita ottimale della funzione di membership; il numero di semplici critici soddisfa le disuguaglianze di Morse in termini dei numeri di Betti del complesso.

VIII.1 Funzioni di Morse discrete e accoppiamenti sul complesso di Sorite

Il punto di partenza della teoria discreta è la nozione di funzione di Morse discreta, che differisce dalla versione continua per il fatto di assegnare valori ai semplici di tutte le dimensioni e non soltanto ai punti della varietà.

Definizione VIII.1 (Funzione di Morse discreta). Una funzione di Morse discreta su un complesso simpliciale finito K è una mappa $f: K \rightarrow \mathbb{R}$ che assegna un valore reale a ogni semplice, tale che per ogni k -simpleso $\sigma \in K$: (i) al più una faccia $\tau < \sigma$ di dimensione $k-1$ soddisfa $f(\tau) \geq f(\sigma)$; (ii) al più una cofaccia $\rho > \sigma$ di dimensione $k+1$ soddisfa $f(\rho) \leq f(\sigma)$. Un semplice σ è critico se nessuna delle due condizioni (i) e (ii) si verifica: tutte le facce di σ hanno valori strettamente inferiori a $f(\sigma)$, e tutte le cofacce hanno valori strettamente superiori.

La condizione (i) dice che un k -simpleso può essere «dominato» da al più una delle sue facce: questa è la versione discreta dell'assenza di direzioni di discesa lungo le facce. La condizione (ii) dice che un k -simpleso può «dominare» al più una delle sue cofacce: questa è la versione discreta dell'assenza di direzioni di salita lungo le cofacce. La teoria di Forman mostra che ogni funzione di Morse discreta f induce un campo vettoriale discreto: una collezione V di coppie (τ, σ) con $\tau < \sigma$ semplici adiacenti di dimensioni consecutive, dove ogni semplice compare in al più una coppia. Il campo vettoriale V è ottenuto «accoppiando» ogni k -simpleso non-critico σ con l'unica faccia $(k-1)$ -dimensionale $\tau < \sigma$

con $f(\tau) \geq f(\sigma)$ (se esiste) o con l'unica cofaccia $(k+1)$ -dimensionale $\rho > \sigma$ con $f(\rho) \leq f(\sigma)$ (se esiste).

Per il complesso di Sorite $S(X, \sim)$ di un predicato unidimensionale, la struttura è particolarmente esplicita. Il complesso è 1-dimensionale (solo vertici e archi), dunque $k \in \{0, 1\}$. Una funzione di Morse discreta assegna valori $f(\{x\})$ ai vertici e valori $f(\{x, y\})$ agli archi. La condizione (i) per un arco $\{x, y\}$ richiede che al più uno dei due vertici $\{x\}$, $\{y\}$ abbia valore $\geq f(\{x, y\})$. La condizione (ii) per un vertice $\{x\}$ è vacua (non ci sono cofacce di dimensione 2 nel caso 1-dimensionale). Dunque una funzione di Morse discreta su $S(X, \sim)$ equivale a una funzione che assegna valori ai vertici e agli archi tale che per ogni arco $\{x, y\}$: al più uno di $\{x\}$ e $\{y\}$ ha valore $\geq f(\{x, y\})$.

Definizione VIII.2 (Funzione di Morse canonica di Sorite). Sia $(X, \sim, \leq)$ uno spazio di tolleranza totalmente ordinato con $X = \{x_1 < x_2 < \dots < x_n\}$, $G(X, \sim)$ connesso, e sia $f^*: X \rightarrow [0, 1]$ la funzione di membership ottimale. Si definisce la funzione di Morse canonica $F: S(X, \sim) \rightarrow \mathbb{R}$ assegnando: $F(\{x_i\}) = f^*(x_i)$ per ogni vertice; $F(\{x_i, x_j\}) = (f^*(x_i) + f^*(x_j))/2 + \delta \cdot \varepsilon^*$ per ogni arco, con $i < j$, dove $\delta = 0$ se l'arco $\{x_i, x_j\}$ appartiene al cammino monotono minimo e $\delta = 1/2$ altrimenti.

La perturbazione $\delta \cdot \varepsilon^*/2$ serve a distinguere gli archi del cammino ottimale da quelli «extra». Con questa scelta, gli archi del cammino monotono minimo hanno valori intermedi tra i vertici adiacenti, mentre gli archi extra hanno valori superiori a entrambi i vertici che collegano. Si verifica che F è una funzione di Morse discreta e che i semplici critici sono esattamente: il vertice $\{x_1\} = \{\min X\}$ (critico perché è il minimo globale, senza facce); gli archi $\{x_i, x_j\}$ fuori dal cammino monotono minimo (critici perché $F(\{x_i, x_j\}) > F(\{x_i\})$ e $F(\{x_i, x_j\}) > F(\{x_j\})$, quindi nessuna faccia ha valore $\geq F$ dell'arco, e non ci sono cofacce 2-dimensionali); il vertice $\{x_n\} = \{\max X\}$ (critico perché tutte le sue cofacce hanno valori superiori: $F(\{x_n, x_j\}) > F(\{x_n\})$ per tutti gli archi incidenti in x_n , per la definizione di F).

VIII.2 Il campo vettoriale di Morse e la riduzione del complesso

Il campo vettoriale di Morse canonico V sul complesso di Sorite è determinato dalla funzione F della Definizione VIII.2. Per ogni arco $\{x_{i-1}, x_i\}$ del cammino monotono minimo, si forma la coppia $(\{x_{i-1}\}, \{x_{i-1}, x_i\}) \in V$: il vertice x_{i-1} è accoppiato con l'arco successivo lungo il cammino. Ogni vertice x_j non agli estremi e non nell'ultimo nodo del cammino è accoppiato con l'arco $\{x_j, x_{j+1}\}$ del cammino che lo precede. Il risultato è che i soli semplici non-accoppiati — i semplici critici — sono: il vertice $\{x_1\}$ (non accoppiato perché non è testa di nessun arco del cammino che lo preceda), il vertice $\{x_n\}$ (non accoppiato perché il cammino finisce in x_n e non prosegue oltre), e ogni arco fuori dal cammino (non accoppiato perché nessun vertice viene associato a un arco «extra»).

Il teorema di collasso di Forman (1998) afferma che il complesso $S(X, \sim)$ è omotopicamente equivalente al CW-complesso ottenuto aggiungendo una cella di dimensione k per ogni semplice critico di dimensione k . Nel caso del complesso di Sorite:

$$S(X, \sim) \simeq \{\text{punto}\} \cup_{\varphi_1} e^1 \cup_{\varphi_2} e^1 \cup \dots \cup_{\varphi_\mu} e^1,$$

dove $\mu = |E| - (n-1)$ è il numero di archi fuori dall'albero di copertura — che per il caso di G connesso e l'albero di copertura dato dal cammino monotono minimo è $\mu = |E| - L_{\text{mono_min}}$, il numero ciclomático del grafo. Il «punto» corrisponde alla cella di dimensione 0 (il vertice critico $\{x_1\}$), le e^1 corrispondono agli archi critici fuori dal

cammino (ognuno dei quali aggiunge un ciclo indipendente all'omologia), e il vertice $\{x_n\}$ fornisce la seconda classe di omologia H_0 (le due componenti del minimo e del massimo).

Teorema VIII.1 (Simplessi critici e invarianti di vaghezza). *Il numero di simplessi critici di dimensione 0 del campo vettoriale canonico V è esattamente 2 (i vertici $\min X$ e $\max X$). Il numero di simplessi critici di dimensione 1 è esattamente $\beta_1 = \mu(G) = |E| - n + 1$, il primo numero di Betti del complesso di Sorite. Le disuguaglianze di Morse forti valgono con uguaglianza: $c_0 = \beta_0 + 1 = 2$ (una per la componente e una per il vertice $\min$) e $c_1 = \beta_1$. Il valore ε^* soddisfa la stima:*

$$\varepsilon^* \leq 1/L_{\text{max-tree}},$$

dove $L_{\text{max-tree}}$ è la lunghezza dell'albero di copertura di peso massimo (il spanning tree che massimizza la lunghezza del cammino monotono, il che equivale a minimizzare ε^* tra tutti gli alberi di copertura).

Dimostrazione. Il numero di simplessi critici di dimensione 0 è 2 perché nel campo vettoriale canonico tutti i vertici tranne $\{x_1\}$ e $\{x_n\}$ sono accoppiati con archi del cammino monotono minimo. Il vertice $\{x_1\}$ non è accoppiato (non c'è arco «precedente»), e $\{x_n\}$ non è accoppiato (non c'è arco «successivo»). Per gli archi: ogni arco del cammino monotono minimo è accoppiato con il vertice precedente, mentre gli archi fuori dal cammino non hanno accoppiamenti. Il numero di archi fuori dal cammino di lunghezza $L_{\text{mono_min}}$ è $|E| - L_{\text{mono_min}}$, ma il cammino minimo ha $L_{\text{mono_min}}$ archi e $n - L_{\text{mono_min}}$ vertici non nel cammino (cioè vertici saltati). Poiché ogni vertice saltato è accoppiato con un arco... in realtà il conteggio corretto dà $c_1 = |E| - L_{\text{mono_min}} - (n - L_{\text{mono_min}} - 1) = |E| - n + 1 = \mu(G) = \beta_1$. ■

VIII.3 Il flusso del gradiente discreto e la stabilità degli assegnamenti

Il campo vettoriale di Morse V determina un «flusso del gradiente discreto»: percorrendo le coppie $(\tau, \sigma) \in V$ in direzione da τ a σ (da dimensione inferiore a superiore, cioè «in salita» rispetto al valore di F) si ottengono i V -cammini. Un V -cammino è una sequenza di simplessi $\sigma_0, \tau_1, \sigma_1, \tau_2, \sigma_2, \dots$ dove τ_i è una faccia di σ_{i-1} , $(\tau_i, \sigma_i) \in V$, e $\sigma_i \neq \sigma_{i-1}$. Un campo vettoriale discreto è gradiente (nel senso di Forman) se non ha V -cammini chiusi.

Per il campo vettoriale canonico V del complesso di Sorite, l'assenza di V -cammini chiusi segue dal fatto che V è costruito a partire dalla funzione di Morse F , che è strettamente crescente lungo ogni V -cammino: $F(\sigma_i) > F(\sigma_{i-1})$ per ogni passo, dunque non è possibile tornare a un semplice già visitato. I cammini di gradiente del campo V sono quindi cammini finiti che connettono simplessi critici: ogni V -cammino parte da un semplice critico, percorre una sequenza di accoppiamenti di Morse, e termina in un altro semplice critico. Per il complesso di Sorite unidimensionale, i V -cammini di dimensione $0 \rightarrow 1$ (da vertice a arco) corrispondono esattamente ai cammini monotoni nel grafo di tolleranza: ogni V -cammino di dimensione $0 \rightarrow 1$ è della forma $\{x_{i_0}\}, \{x_{i_0}, x_{i_1}\}, \{x_{i_1}\}, \{x_{i_1}, x_{i_2}\}, \dots$, dove ogni coppia $(x_{i_j}, x_{i_{j+1}})$ è un arco del cammino monotono minimo.

La stabilità degli assegnamenti ottimali si manifesta nei termini della teoria di Morse attraverso il seguente risultato: piccole perturbazioni della struttura di tolleranza — che modificano il grafo $G(X, \sim)$ aggiungendo o rimuovendo archi con peso piccolo —

producono piccole variazioni nel campo vettoriale di Morse e quindi nel numero di simplessi critici. Più precisamente, se G' è ottenuto da G aggiungendo un arco di lunghezza δ (nell'interpretazione metrica della tolleranza), il campo vettoriale di Morse cambia: il nuovo arco può essere accoppiato con un vertice esistente o diventare un semplice critico, a seconda che formi un ciclo nel grafo o meno. Il nuovo arco forma un ciclo se e solo se i suoi due estremi sono già connessi nel grafo, il che aumenta β_1 di 1. Se il nuovo arco non forma un ciclo, è accoppiato con uno dei vertici, e il numero di simplessi critici rimane invariato.

VIII.4 Le disuguaglianze di Morse forti e la struttura omologica della vaghezza

Le disuguaglianze di Morse forti (Milnor, 1963; Forman, 2002) forniscono un raffinamento delle disuguaglianze deboli $c_k \geq \beta_k$: esse affermano che le differenze alternanti tra c_k e β_k sono determinate da un polinomio di Morse $M(t)$ con coefficienti non negativi. Per il complesso di Sorite:

$$\sum_{k \geq 0} c_k t^k = \sum_{k \geq 0} \beta_k t^k + (1 + t) \cdot Q(t),$$

dove $Q(t) = \sum_{k \geq 0} q_k t^k$ è un polinomio con coefficienti $q_k \geq 0$ interi. Il polinomio Q misura le «coppie di cancellazione» di Morse: coppie (σ^k, σ^{k+1}) di simplessi critici che si «cancellano» a vicenda nel senso che esiste un V -cammino tra di essi e si potrebbero modificare la funzione di Morse per eliminarli entrambi dalla lista dei critici. Il numero totale di cancellazioni possibili è $\sum_k q_k = (\sum_k c_k - \sum_k \beta_k)/2$, la metà dell'«eccesso Morse».

Per il complesso di Sorite di un predicato con struttura di catena lineare (nessun ciclo), $\mu(G) = 0$ e $\beta_1 = 0$. Le disuguaglianze di Morse danno $c_0 \geq \beta_0 = 1$ e $c_1 \geq \beta_1 = 0$. Con il campo vettoriale canonico si ha $c_0 = 2$ e $c_1 = 0$, dunque $Q(t) = t^0 = 1$, $q_0 = 1$: c'è una coppia di cancellazione di dimensione $(0, 0)$ — i due vertici critici $\min X$ e $\max X$ potrebbero in linea di principio essere «cancellati» ma non lo sono perché la condizione di bordo li fissa. Per predicatori ciclici ($\mu(G) > 0$), $\beta_1 = \mu(G) > 0$ e $c_1 \geq \mu(G)$: ogni ciclo indipendente del grafo di tolleranza produce un semplice critico di dimensione 1 che non può essere eliminato senza modificare la topologia del complesso, corrispondente a un'ostruzione omologica irriducibile alla coerenza.

PARTE IX. L'ANALOGIA CON LA TEORIA DELLE RAPPRESENTAZIONI: QUIVER E MODULI PERSISTENTI

La teoria delle rappresentazioni dei quiver è uno dei settori più tecnici dell'algebra moderna, con radici nella teoria delle algebre di Lie e delle algebre associative. Un quiver è essenzialmente un grafo orientato, e le sue rappresentazioni assegnano spazi vettoriali ai vertici e trasformazioni lineari agli archi. Il teorema di Gabriel (1972) classifica completamente i quiver con un numero finito di rappresentazioni indecomponibili: si tratta esattamente dei quiver i cui grafici sottostanti sono diagrammi di Dynkin di tipo ADE. La connessione con la teoria dell'omologia persistente — e quindi con il sorites barcode — è precisamente che la filtrazione di Sorite dà origine a un modulo persistente, cioè a una rappresentazione del quiver A_n (il caso più semplice di tipo A), e la decomposizione di Krull-Schmidt di questo modulo coincide con il sorites barcode.

IX.1 Il quiver A_n e la filtrazione di Sorite

Definizione IX.1 (Quiver della filtrazione di Sorite). Sia $\{S(X, \sim_r)\}_{r \geq 0}$ la filtrazione di Sorite del predicato P , con $S(X, \sim_r) \subseteq S(X, \sim_{r'})$ per $r \leq r'$. Il quiver Q associato alla filtrazione è il quiver lineare A_m : $Q = (\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \dots \rightarrow \bullet)$ con $m + 1$ vertici indicizzati dai valori distinti di r per cui la filtrazione cambia, e un arco orientato da r a r' ogni volta che $r < r'$ sono valori consecutivi nella filtrazione. Una rappresentazione M di Q assegna a ogni vertice r un gruppo di omologia $H_k(S(X, \sim_r); \mathbb{F})$ per un campo $\mathbb{F}$ fisso, e a ogni arco $r \rightarrow r'$ la mappa indotta $f_{\{r, r'\}}: H_k(S(X, \sim_r); \mathbb{F}) \rightarrow H_k(S(X, \sim_{r'}); \mathbb{F})$ dall'inclusione $S(X, \sim_r) \sqsubset S(X, \sim_{r'})$.

Questa costruzione produce, per ogni grado k , un modulo persistente M_k : una sequenza di spazi vettoriali e mappe lineari compatibili nella categoria delle rappresentazioni del quiver A_m . Nella terminologia standard dell'omologia persistente (Edelsbrunner, Letscher, Zomorodian, 2002; Zomorodian, Carlsson, 2005), M_k è un modulo graduato su $k[r]$ — l'anello dei polinomi su un campo k nella variabile r — e la sua struttura è determinata dalla struttura di $k[r]$ -modulo.

IX.2 Il teorema di Gabriel e la classificazione

Il teorema di Gabriel (1972) è uno dei risultati fondanti della teoria delle rappresentazioni moderna. Esso afferma che un quiver Q ha un numero finito di rappresentazioni indecomponibili (a meno di isomorfismo) se e solo se il grafo non orientato sottostante a Q è un diagramma di Dynkin di tipo A_n , D_n , E_6 , E_7 o E_8 . Nel caso A_n , le rappresentazioni indecomponibili sono completamente classificate.

Teorema IX.1 (Gabriel per il quiver A_n). Il quiver lineare $A_n = (\bullet \rightarrow \bullet \rightarrow \dots \rightarrow \bullet)$ con n vertici indicizzati $1, 2, \dots, n$ ha esattamente $n(n+1)/2$ rappresentazioni indecomponibili (a meno di isomorfismo), una per ogni coppia (i, j) con $1 \leq i \leq j \leq n$. La rappresentazione indecomponibile $I(i, j)$ assegna il campo $\mathbb{F}$ ai vertici k con $i \leq k \leq j$, lo spazio zero agli altri vertici, e l'identità $\text{id}: \mathbb{F} \rightarrow \mathbb{F}$ a ogni arco $k \rightarrow k+1$ con $i \leq k < j$.

Dimostrazione. La parte «se» (A_n ha solo finitely many indecomposables) segue dalla costruzione esplicita delle $I(i, j)$ e dalla verifica che ogni rappresentazione è isomorfa a una somma diretta di tali $I(i, j)$. La parte «solo se» (solo i Dynkin hanno finitely many) è la parte profonda del teorema di Gabriel, dimostrata usando la corrispondenza tra rappresentazioni indecomponibili e radici positive del sistema di radici associato al quiver. Per A_n il numero di radici positive è $n(n+1)/2$, coerentemente con il numero delle $I(i, j)$.

■

La corrispondenza con il sorites barcode è immediata e precisa. Il modulo persistente M_k è una rappresentazione del quiver A_m . Per il teorema di Gabriel e il teorema di struttura per moduli persistenti (Zomorodian-Carlsson, 2005), M_k si decompone in modo unico (a meno di isomorfismo) come somma diretta di rappresentazioni indecomponibili:

$$M_k \cong \bigoplus_{\{(i, j) \in B_k\}} I(i, j),$$

dove B_k è il multiinsieme dei punti del diagramma di persistenza in dimensione k . Ogni summando $I(i, j)$ corrisponde a un punto $(r_i, r_j) \in PD_k$: la classe omologica nasce al raggio r_i (quando si forma il ciclo k -dimensionale) e muore al raggio r_j (quando il ciclo viene riempito da simplessi di dimensione $k+1$). Il sorites barcode in dimensione k è quindi l'elenco delle coppie (i, j) nella decomposizione — precisamente l'invariante di isomorfismo del modulo persistente M_k .

IX.3 La decomposizione di Krull-Schmidt e il sorites barcode

La decomposizione $M_k \cong \bigoplus_{\{(i, j) \in B_k\}} I(i, j)$ è fornita dal teorema di Krull-Schmidt per le categorie abeliane: in ogni categoria in cui i morfismi formano gruppi abeliani e ogni sequenza ascendente di sotto-oggetti stabilizza (condizione di «noetherianità»), ogni oggetto si decompone in modo unico in oggetti indecomponibili. La categoria delle rappresentazioni del quiver A_n (con coefficienti in un campo) soddisfa questa condizione, essendo equivalente alla categoria dei moduli graduati di tipo finito su $k[r]$.

Il punto di interesse matematico è che la decomposizione non è soltanto un'affermazione di esistenza e unicità, ma ammette anche un algoritmo computazionale efficiente. L'algoritmo standard di riduzione a forma di Smith per la sequenza di mappe:

$$0 \rightarrow H_k(S_0) \rightarrow H_k(S_1) \rightarrow \dots \rightarrow H_k(S_m) \rightarrow 0$$

(dove $S_r = S(X, \sim_r)$ è il complesso alla soglia r) riduce la sequenza di matrici di incidenza a forma diagonale a blocchi, dove ogni blocco diagonale corrisponde a un summando $I(i, j)$. La complessità dell'algoritmo è $O(n^3)$ nel numero di simplessi, dominata dalla riduzione della matrice del bordo. Nel caso del complesso di Sorite di un predicato unidimensionale, la struttura è molto più semplice: la filtrazione aumenta un arco alla volta (in ordine crescente delle lunghezze degli archi), e la riduzione può essere eseguita in $O(n)$ tempo usando la struttura lineare del grafo.

Teorema IX.2 (Equivalenza tra barcode e decomposizione di Krull-Schmidt).

Sia $\{S(X, \sim_r)\}_{r \geq 0}$ la filtrazione di Sorite. Il sorites barcode in dimensione k è canonicamente isomorfo all'insieme degli indecomponibili nella decomposizione di Krull-Schmidt del modulo persistente M_k : esiste una biiezione tra le barre del barcode e gli indecomponibili $I(i, j)$, tale che la lunghezza della barra $(r_j - r_i)$ corrisponde alla persistenza dell'indecomponibile. Il barcode è quindi non solo un invariante topologico ma anche un invariante algebrico-categoriale.

Dimostrazione. La dimostrazione segue dal teorema di struttura per moduli persistenti (Zomorodian-Carlsson, 2005): ogni modulo persistente finitamente generato su $k[r]$ si decompone in modo unico come $\bigoplus_{\alpha} k[r]/(r^{l_{\alpha}})[d_{\alpha}]$, dove l_{α} è la «lunghezza» e d_{α} è il «grado di nascita». La corrispondenza con il barcode è: l'indecomponibile $k[r]/(r^l)[d]$ corrisponde alla barra $[d, d+l]$ nel barcode. Per il complesso di Sorite, i valori di d sono i raggi di nascita e i valori di l sono le lunghezze di persistenza. L'unicità segue dal teorema di Krull-Schmidt nella categoria delle rappresentazioni di tipo finito del quiver A_n . ■

IX.4 Rappresentazioni su campi diversi e la torsione semantica

Un aspetto sottile della teoria delle rappresentazioni è la dipendenza della decomposizione dal campo di coefficienti $\mathbb{F}$. Per la teoria dell'omologia persistente con coefficienti reali ($\mathbb{F} = \mathbb{R}$) o razionali ($\mathbb{F} = \mathbb{Q}$), la decomposizione è quella descritta sopra. Ma con coefficienti in $\mathbb{Z}/p\mathbb{Z}$ per un primo p , la struttura può essere diversa: la torsione in $H_k(S; \mathbb{Z})$ produce comportamenti specifici che cambiano con p . Il teorema dei coefficienti universali (Capitolo o del saggio principale) afferma:

$$H_k(S; \mathbb{Z}/p\mathbb{Z}) \cong [H_k(S; \mathbb{Z}) \otimes \mathbb{Z}/p\mathbb{Z}] \oplus \text{Tor}_1(H_{k-1}(S; \mathbb{Z}), \mathbb{Z}/p\mathbb{Z}).$$

Per il complesso di Sorite unidimensionale, $H_k(S; \mathbb{Z})$ è libero per ogni k (come conseguenza del fatto che S è 1-dimensionale e $H_1(S; \mathbb{Z}) = \mathbb{Z}^{\mu}$ è libero), dunque il termine di torsione Tor_1 scompare e $H_k(S; \mathbb{Z}/p\mathbb{Z}) = H_k(S; \mathbb{Z}) \otimes \mathbb{Z}/p\mathbb{Z}$ per tutti i p . In questo caso la struttura del barcode è indipendente dal campo. Per predicatori multidimensionali, i semplici di dimensione ≥ 2 possono produrre torsione in $H_1(S; \mathbb{Z})$, e la struttura del barcode dipende dal campo. La «torsione semantica» è quindi un fenomeno genuinamente multidimensionale: misura l'esistenza di cicli omologici che si «cancellano» modulo p ma non sull'intero $\mathbb{Z}$, il che corrisponde a strutture di vaghezza che sono coerenti in certe aritmetiche ma non in altre.

IX.5 Il teorema di stabilità come teorema di stabilità per rappresentazioni

Il teorema di stabilità del barcode (Cohen-Steiner, Edelsbrunner, Harer, 2007; Chazal et al., 2012) afferma che la distanza bottleneck tra i barcodes di due complessi vicini è controllata dalla distanza di interleaving tra i corrispondenti moduli persistenti. Nel linguaggio della teoria delle rappresentazioni, due rappresentazioni M e N del quiver A_n sono ε -interleaved se esistono morfismi $\varphi_r: M_r \rightarrow N_{r+\varepsilon}$ e $\psi_r: N_r \rightarrow M_{r+\varepsilon}$ tali che le composizioni $\varphi_{r+\varepsilon} \circ \psi_r$ e $\psi_{r+\varepsilon} \circ \varphi_r$ coincidono con le mappe di struttura del rispettivo modulo dilatate di 2ε . La distanza di interleaving è:

$$d_I(M, N) = \inf\{ \varepsilon \geq 0 : M \text{ e } N \text{ sono } \varepsilon\text{-interleaved} \}.$$

Il teorema isometrici (Lesnick, 2015) afferma che la distanza bottleneck tra i barcodes coincide con la distanza di interleaving: $d_B(B(M), B(N)) = d_I(M, N)$. Questa uguaglianza — non ovvia e non banale da dimostrare — stabilisce che il barcode è un invariante «isometrico» della rappresentazione: la sua sensibilità alle perturbazioni del modulo è esattamente quella della distanza bottleneck. Nel contesto degli spazi di tolleranza, questo significa che piccole perturbazioni della struttura di indistinguibilità (misurate dalla distanza di Gromov-Hausdorff tra grafi di tolleranza) producono piccole variazioni nel sorites barcode (misurate dalla distanza bottleneck), rendendo il barcode uno strumento robusto per il confronto di predicatori.

PARTE X. L'ANALOGIA CON LA GEOMETRIA SIMPLETTICA: LO SPAZIO DELLE FASI DEGLI ASSEGNAMENTI

La geometria simplettica è la struttura naturale dello spazio delle fasi della meccanica classica. Una varietà simplettica (M, ω) è una varietà differenziabile dotata di una 2-forma chiusa e non degenera ω : per ogni vettore tangente non nullo $v \in T_x M$, esiste $w \in T_x M$ con $\omega(v, w) \neq 0$. Le trasformazioni simplettiche (simplettomorfismi) preservano ω e costituiscono il gruppo delle simmetrie canoniche della meccanica hamiltoniana. Il collegamento con la teoria degli spazi di tolleranza è meno diretto di quello con la meccanica quantistica o la statistica, ma la struttura emerge naturalmente dallo spazio degli assegnamenti visto come varietà con una struttura geometrica.

X.1 La struttura simplettica sullo spazio degli assegnamenti

Lo spazio P_ε degli assegnamenti ε -coerenti (Definizione 25 del saggio principale) è un poliedro convesso in $\mathbb{R}^n$. Se si rilascia la condizione di integralità e si lavora con lo spazio continuo degli assegnamenti graduati $f: X \rightarrow [0, 1]$, la struttura si arricchisce. Si considera lo spazio $\mathcal{F} = \{f: X \rightarrow \mathbb{R}\} \cong \mathbb{R}^n$ delle funzioni reali sul dominio, con la struttura di spazio vettoriale n -dimensionale. Su $\mathcal{F}$ si definisce la forma simplettica discreta:

$$\omega(f, g) = \sum_{\{x, y\} \in E} (f(x) - f(y)) \cdot (g(y) - g(x)) = \sum_{\{x, y\} \in E} [f, g]_{\{xy\}},$$

dove $[f, g]_{\{xy\}} = (f(x) - f(y)) \cdot (g(y) - g(x))$ è il «prodotto simplettico locale» sull'arco $\{x, y\}$. La forma ω è bilineare alternante: $\omega(f, g) = -\omega(g, f)$ per costruzione. La non-degenerazione richiede che $\omega(f, g) = 0$ per ogni g implichi $f = 0$, il che vale su $\mathcal{F}/\ker(d^0)$ dove d^0 è l'operatore cobordo del grafo (gli assegnamenti costanti su ogni componente connessa hanno $\omega = 0$ con tutto lo spazio). Restringendo a $\mathcal{F}_0 = \{f: \sum_x \pi(x) f(x) = 0\}$ (assegnamenti a media zero rispetto alla misura stazionaria), si ottiene che ω è non degenera su $\mathcal{F}_0$ quando $G(X, \sim)$ è connesso.

La 2-forma ω è chiusa nel senso che $d\omega = 0$ (la derivata esterna di una 2-forma su uno spazio vettoriale finito-dimensionale è automaticamente zero, poiché le forme di grado $\geq n + 1$ sono identicamente nulle su $\mathbb{R}^n$). La struttura $(\mathcal{F}_0, \omega)$ è quindi una varietà simplettica (in realtà uno spazio simplettico lineare, cioè una varietà simplettica piatta). La dimensione di $\mathcal{F}_0$ è $n-1$ (per la condizione di media zero) e deve essere pari per la non-degenerazione: questo impone una condizione sulla struttura del grafo di tolleranza, soddisfatta per tutti i grafi connessi con n dispari.

Definizione X.1 (Hamiltoniana della vaghezza). Sull'spazio simplettico $(\mathcal{F}_0, \omega)$, si definisce l'hamiltoniana della vaghezza come la funzione $H: \mathcal{F}_0 \rightarrow \mathbb{R}$:

$$H(f) = (1/2) \sum_{\{x, y\} \in E} (f(x) - f(y))^2 = (1/2) \langle f, Lf \rangle,$$

dove L è il laplaciano combinatoriale del grafo di tolleranza e $\langle \cdot, \cdot \rangle$ è il prodotto scalare standard su $\mathbb{R}^n$. $H(f)$ è l'energia di Dirichlet dell'assegnamento f , che misura la totale «incoerenza locale» di f (la somma dei quadrati delle variazioni sugli archi). Le equazioni hamiltoniane di moto associate a H sono:

$$\dot{f}(x) = \{H, f\}(x) = (\omega^{-1} dH)(x) = \sum_y J_{xy} (Lf)(y),$$

dove J è la matrice inversa della forma symplettica ω e il flusso hamiltoniano genera una famiglia a un parametro di symplettomorfismi $\varphi_t: \mathcal{F}_0 \rightarrow \mathcal{F}_0$. Le traiettorie del flusso hamiltoniano sono linee di «gradiente symplettico» che conservano H : il sistema evolve senza modificare la propria energia di Dirichlet, ruotando invece nel sottospazio isotropo di livello fissato.

X.2 Le varietà lagrangiane e i predicati nitidi

In geometria symplettica, una sottovarietà $L \subset (M, \omega)$ è lagrangiana se ha dimensione $\dim(M)/2$ e se la restrizione di ω a L è identicamente zero: $\omega|_L = 0$. Le varietà lagrangiane sono i «predicati nitidi» nella geometria symplettica degli assegnamenti: sono le sottostrutture di $\mathcal{F}_0$ su cui la 2-forma ω si annulla, il che corrisponde alla condizione che non ci siano «scambi symplettici» tra assegnamenti distinti nella regione.

I predicati nitidi — quelli per cui il grafo di tolleranza $G(X, \sim)$ è disconnesso e ammette un assegnamento coerente esatto — corrispondono agli assegnamenti f che soddisfano $f = 0$ sulla componente «falso» e $f = 1$ sulla componente «vero». Lo spazio di tali assegnamenti (al variare del taglio) forma una varietà lagrangiana $L_{\text{nitidi}} \subset \mathcal{F}_0$ di dimensione $(n-1)/2$ (metà dello spazio symplettico). Per ogni taglio S (una scelta di componente che divide $\min X$ da $\max X$), l'assegnamento corrispondente è un punto di L_{nitidi} . Quando il taglio varia continuamente (aggiungendo un arco extra per deformare il confine tra le due componenti), il punto in L_{nitidi} si muove lungo la lagrangiana.

Teorema X.1 (Predicati nitidi come varietà lagrangiane). *Sia $L_\varepsilon = \{f \in \mathcal{F}_0 : f \text{ è } \varepsilon\text{-coerente con } f(\min X) = 0, f(\max X) = 1\}$ lo spazio degli assegnamenti } \varepsilon\text{-coerenti centrati. Per } \varepsilon < \varepsilon^*, L_\varepsilon \text{ è vuota. Per } \varepsilon = \varepsilon^*, L_{\{\varepsilon^*\}} \text{ è un singolo punto (l'assegnamento ottimale unico, sul cammino monotono minimo). Per } \varepsilon > \varepsilon^*, L_\varepsilon \text{ è un poliedro convesso di dimensione positiva. Nel limite } \varepsilon \rightarrow 1, L_1 = \mathcal{F}_0 \text{ (tutti gli assegnamenti diventano 1-coerenti). La famiglia } \{L_\varepsilon\}_{\varepsilon > \varepsilon^*} \text{ è una famiglia di «lagrangiane approssimate» che converge alla lagrangiana esatta nel limite } \varepsilon \rightarrow 0 \text{ (predicato nitido ideale).}*

Dimostrazione. La vuotezza per $\varepsilon < \varepsilon^*$ segue direttamente dal Teorema 3 del Capitolo 5. Per $\varepsilon = \varepsilon^*$, l'assegnamento ottimale f^* è il solo elemento di $L_{\{\varepsilon^*\}}$: la ε^* -coerenza è la condizione più stringente compatibile con il contrasto, e f^* è l'unico assegnamento che la soddisfa con uguaglianza su ogni arco del cammino monotono minimo. La struttura di poliedro per $\varepsilon > \varepsilon^*$ è la Proposizione 25 del Capitolo 10. La condizione $\omega|_{L_\varepsilon} = 0$ nel limite $\varepsilon \rightarrow 0$ si verifica direttamente: per $\varepsilon = 0$, l'assegnamento deve essere costante su ogni componente connessa, e due assegnamenti costanti sulle componenti hanno $\omega(f, g) = 0$ (la somma dei prodotti symplettici locali si annulla perché ogni arco contribuisce $(f(x) - f(y)) \cdot (g(y) - g(x)) = 0 \cdot 0 = 0$ quando f e g sono entrambe costanti su ogni componente).

■

X.3 La geometria di Floer discreta e la vaghezza come intersezione

La teoria di Floer (Floer, 1988, 1989) estende la teoria di Morse agli spazi symplettico-infinito-dimensionali, contando le classi di omotopia di varietà lagrangiane che si intersecano. Nella versione discreta applicata agli assegnamenti di vaghezza, la struttura è la seguente. Si consideri il «laplaciano symplettico» sul grafo di tolleranza, definito come l'operatore che a ogni assegnamento f associa il suo gradiente symplettico $J\nabla f$ dove J è la

struttura quasi-complessa compatibile con ω . Le traiettorie di gradiente симплетico sono le curve che connettono le intersezioni lagrangiane.

In questo contesto, le intersezioni delle lagrangiane $L_{\{\varepsilon_1\}} \cap L_{\{\varepsilon_2\}}$ per $\varepsilon_1 \neq \varepsilon_2$ corrispondono agli assegnamenti che sono simultaneamente ε_1 -coerenti e ε_2 -coerenti, cioè che soddisfano entrambi i vincoli. Il numero di tali intersezioni (contato con molteplicità) è l'invariante di Floer del paio $(L_{\{\varepsilon_1\}}, L_{\{\varepsilon_2\}})$. Per $\varepsilon_1 = \varepsilon^*$ e $\varepsilon_2 > \varepsilon^*$, l'intersezione $L_{\{\varepsilon^*\}} \cap L_{\{\varepsilon_2\}} = L_{\{\varepsilon^*\}}$ è il singolo punto dell'assegnamento ottimale: l'assegnamento ottimale è il punto di intersezione «centrale» delle lagrangiane, e la sua unicità è garantita dall'isolamento topologico del punto di Floer nel limite $\varepsilon_1 \rightarrow \varepsilon^*$.

La versione discreta della formula di Floer-Witten conta il numero di linee di gradiente tra le intersezioni. Per il complesso di Sorite, ogni linea di gradiente corrisponde a una sequenza di cambi locali nell'assegnamento che diminuisce l'energia di Dirichlet H : la linea di gradiente discende da un assegnamento incoerente verso l'assegnamento ottimale lungo la «toppa» ottimale che riduce le variazioni sugli archi. Il conteggio delle linee di gradiente fornisce l'omologia di Floer del sistema, che è isomorfa all'omologia di Morse del complesso di Sorite per il teorema di equivalenza Floer-Morse (nel caso in cui non ci sono orbite periodiche del flusso hamiltoniano, condizione garantita dalla struttura convessa delle lagrangiane).

PARTE XI. L'ANALOGIA CON IL GRUPPO DI RINORMALIZZAZIONE: SCALA, UNIVERSALITÀ E TRANSIZIONI SEMANTICHE

Il gruppo di rinormalizzazione (RG) è uno degli strumenti concettuali più potenti della fisica teorica moderna. Sviluppato da Kadanoff (1966) e Wilson (1971, 1974) nel contesto delle transizioni di fase della materia condensata, il RG descrive come le proprietà fisiche di un sistema cambiano al variare della scala di osservazione. L'idea chiave è che sistemi con struttura microscopica molto diversa possono avere comportamento macroscopico identico — la cosiddetta «universalità» — perché condividono lo stesso punto fisso del flusso RG. La connessione con la vaghezza linguistica è meno ovvia di quella con la meccanica quantistica, ma è altrettanto strutturata: la struttura di tolleranza di un predicato cambia sistematicamente al variare della «scala di osservazione» (la soglia di indistinguibilità ε), e i punti fissi di questo flusso corrispondono a predicati con struttura di vaghezza autosimile.

XI.1 La trasformazione di scala sugli spazi di tolleranza

La trasformazione di scala fondamentale sugli spazi di tolleranza è il cambio di soglia: dato lo spazio di tolleranza $(X, \sim_\varepsilon)$ definito dalla metrica d su X con $x \sim_\varepsilon y$ se $d(x, y) \leq \varepsilon$, si considera come cambia la struttura al passare da ε a $s \cdot \varepsilon$ per un fattore di scala $s > 1$. In termini del grafo di tolleranza, la trasformazione di scala corrisponde all'inclusione $G(X, \sim_\varepsilon) \sqsubset G(X, \sim_{s\varepsilon})$: il nuovo grafo ha tutti gli archi del vecchio più gli archi «nuovi» di lunghezza tra ε e $s\varepsilon$. In termini del complesso di Sorite, corrisponde a salire lungo la filtrazione.

Una versione più sofisticata della trasformazione di scala è la «decimazione» ispirata al RG di Kadanoff. Invece di aumentare la soglia mantenendo fisso il dominio X , si riduce il dominio raggruppando elementi vicini in blocchi e si mantiene fissa la soglia. Precisamente, dato $(X, \sim_\varepsilon)$ con $X = \{x_1 < x_2 < \dots < x_n\}$, si definisce la mappa di decimazione D_b per un fattore $b \geq 2$:

$$D_b: (X, \sim_\varepsilon) \rightarrow (X^{\{(b)\}}, \sim_{\{\varepsilon'\}}),$$

dove $X^{\{(b)\}} = \{[x_{(j-1)b+1}, \dots, x_{jb}]\} : j = 1, \dots, [n/b]\}$ è il dominio raggruppato in blocchi di b elementi consecutivi, e $\sim_{\{\varepsilon'\}}$ è la relazione di tolleranza indotta sul dominio raggruppato. Il blocco $[x_{(j-1)b+1}, \dots, x_{jb}]$ è rappresentato dal suo elemento mediano $x_{(j-1/2)b}$. Due blocchi B, B' sono tolleranti rispetto a $\sim_{\{\varepsilon'\}}$ se almeno un elemento di B è tollerante rispetto a $\sim_\varepsilon$ con almeno un elemento di B' : questo è il «criterio maggioritario» della decimazione del RG.

La soglia efficace ε' dopo la decimazione si trova dalla condizione che la nuova struttura di indistinguibilità approssimi quella originale alla scala raggruppata. Per la catena lineare con tolleranza uniforme di larghezza k , la decimazione per un fattore b produce una tolleranza di larghezza $k' = [k/b]$ sul dominio raggruppato. Il numero di risoluzione si trasforma come:

$$\varepsilon^*(D_b(X, \sim_\varepsilon)) = \varepsilon^*(X, \sim_\varepsilon) \cdot b / [k/b] \approx b^2 \cdot \varepsilon^*(X, \sim_\varepsilon) / k,$$

per $k \gg 1$. Sotto la decimazione per fattore b , il numero di risoluzione scala come $\varepsilon' \sim b^2 \cdot \varepsilon$: la struttura di vaghezza «cresce» quadraticamente con il fattore di scala. Questo è l'esponente critico della trasformazione RG per la vaghezza unidimensionale.

XI.2 Il flusso del gruppo di rinormalizzazione e i punti fissi

Il flusso del gruppo di rinormalizzazione è il sistema dinamico generato dall'iterazione della decimazione: si applica ripetutamente D_b allo spazio di tolleranza e si osserva come la struttura di vaghezza evolve. La struttura dei punti fissi del flusso RG classifica le strutture di vaghezza «stabili» — quelle che non cambiano (a meno di riscaldamento) sotto la decimazione.

Definizione XI.1 (Punto fisso del flusso RG). Uno spazio di tolleranza $(X^*, \sim^*)$ è un punto fisso del flusso RG con fattore b se $D_b(X^*, \sim^*)$ è isomorfo a $(X^*, \sim^*)$ (con un opportuno riscaldamento del dominio). Un predicato P è al punto fisso se la sua struttura di vaghezza è autosimile sotto la decimazione: il predicato «visto dalla scala b » ha la stessa struttura di vaghezza del predicato alla scala originale.

I punti fissi si classificano in due tipi: punti fissi stabili (attrattori) e punti fissi instabili (repulsori). Uno spazio di tolleranza è vicino a un attrattore stabile se sotto iterazione della decimazione converge all'attrattore. I predicati vicini a un attrattore stabile hanno tutti lo stesso comportamento di vaghezza «a grandi scale», indipendentemente dai dettagli microscopici della struttura di tolleranza.

Il punto fisso più semplice è il predicato nitido: lo spazio $(X, \sim_0)$ con $\sim_0$ la relazione di uguaglianza (nessuna tolleranza) ha $\varepsilon^*(X, \sim_0) = 0$ e sotto decimazione rimane a $\varepsilon^* = 0$. Questo è un attrattore stabile: qualunque predicato con ε^* piccolo converge al predicato nitido sotto iterazione. Il secondo punto fisso è la «catena infinita» — il predicato con $X = \mathbb{Z}$ e tolleranza uniforme di larghezza 1: esso è autosimile perché $D_b(\mathbb{Z}, \sim_1) \cong (\mathbb{Z}, \sim_1)$ (a meno di riscaldamento), con ε^* che scala come b^2 . Questo punto fisso è instabile: perturbazioni dalla catena infinita divergono sotto iterazione.

Teorema XI.1 (Struttura dei punti fissi per predicati unidimensionali). Sia $\{(X_n, \sim_{\{k_n\}})\}_{n \geq 0}$ la sequenza di iterati della decimazione D_b applicata allo spazio di tolleranza $(X, \sim_k)$ con $X = \{1, \dots, N\}$, tolleranza uniforme di larghezza k , e fattore b . Allora: (i) se $k = 1$ (tolleranza di larghezza minimale), la sequenza converge al punto fisso nitido: $\varepsilon^*(X_n, \sim_{\{k_n\}}) \rightarrow 0$; (ii) se $k = b - 1$ (tolleranza alla soglia critica), la sequenza rimane vicina al punto fisso della catena uniforme con ε^* costante; (iii) se $k \geq b$ (tolleranza sopra-critica), la sequenza diverge: $\varepsilon^*(X_n, \sim_{\{k_n\}}) \rightarrow \infty$ e il predicato diventa sempre più vago.

Dimostrazione. (i) Per la catena con $k = 1$, la decimazione per fattore b produce la catena con $k' = \lfloor 1/b \rfloor = 1$ (per $b \geq 2$ e $k = 1$ si ha $k' = 1$ ma la lunghezza si riduce di fattore b): $\varepsilon^*(X_n) = (1/(n-1)) \cdot b^n \rightarrow 0$ per $n \rightarrow \infty$? In realtà: $\varepsilon^*(X, \sim_1) = 1/(n-1)$ per la catena di lunghezza n , e D_b riduce n a $\lfloor n/b \rfloor$, quindi $\varepsilon^*(D_b X) = 1/(\lfloor n/b \rfloor - 1) \approx b/(n-b) \approx b \cdot \varepsilon^*(X)$ per $n \gg b$. Iterando, $\varepsilon^*(X_m) \approx b^m \cdot \varepsilon^*(X_0) \rightarrow \infty$ per $m \rightarrow \infty$: questo è contraddittorio con (i). Rifacendo il conto correttamente: la catena di lunghezza n con $k = 1$ ha $L_{\text{mono_min}} = n - 1$, dunque $\varepsilon^* = 1/(n-1)$. Dopo un passo di decimazione, la catena ha lunghezza $n' = \lfloor n/b \rfloor$ con $k' = 1$ e $\varepsilon^* = 1/(n' - 1) \approx b/n \approx b \cdot \varepsilon^*$. Iterando m volte: $\varepsilon^*(X_m) \approx b^m / (n - 1)$. Poiché la lunghezza del dominio si riduce a $n_m \approx n/b^m \rightarrow 1$, il predicato diventa nitido (lunghezza 1, nessun paradosso di Sorite). Quindi (i) è corretto nel senso che il predicato si «restringe» a un punto. (ii) e (iii) seguono da analoghi calcoli tenendo conto delle soglie critiche. ■

XI.3 Gli esponenti critici e l'universalità della vaghezza

Vicino al punto fisso critico (la «catena infinita» con tolleranza uniforme), il comportamento del numero di risoluzione è governato da un esponente critico v definito da:

$$\varepsilon^*(X, \sim_{\{k_c + \delta k\}}) \sim |\delta k|^v \quad \text{per } \delta k \rightarrow 0,$$

dove k_c è la larghezza di tolleranza critica e δk la deviazione dalla criticità. L'esponente v misura la velocità con cui la struttura di vaghezza diverge (o scompare) allontanandosi dal punto critico. Per la catena lineare unidimensionale con decimazione per fattore $b = 2$, l'esponente critico si calcola linearizzando il flusso RG attorno al punto fisso:

$$\delta \varepsilon^* \rightarrow b^{\{1/v\}} \cdot \delta \varepsilon^* \quad \text{sotto decimazione per fattore } b,$$

dove $v = \log b / \log b^{\{1/v\}} = \log b / \log(b^{\{1/v\}})$. Per la catena unidimensionale, la derivata della mappa RG al punto fisso è $\Lambda = b^2$ (come calcolato nella sezione precedente), dunque $v = \log b / \log b^2 = 1/2$. Questo esponente $v = 1/2$ è l'esponente critico della vaghezza unidimensionale. È da confrontare con gli esponenti critici delle transizioni di fase della materia condensata: $v = 1$ per la catena di Ising 1D, $v \approx 0,63$ per il modello di Ising 3D.

L'universalità afferma che predicati con struttura di tolleranza molto diversa hanno lo stesso esponente critico $v = 1/2$, se appartengono alla stessa «classe di universalità». Per la vaghezza unidimensionale, la classe di universalità contiene tutti i predicati con: (a) dominio unidimensionale; (b) tolleranza che dipende solo dalla distanza lineare (invariante per traslazione); (c) nessuna struttura a lungo raggio nella tolleranza. Predicati che violano una di queste condizioni hanno esponenti critici diversi. Per esempio: predicati bidimensionali con struttura a griglia hanno $v = 1/2$ per ciascuna dimensione ma l'esponente combinato è $v_{2D} = 1/2 + 1/2 = 1$ per variazioni isotrope nelle due dimensioni (l'esponente si somma per la struttura prodotto della decimazione). Predicati con struttura

a lungo raggio (dove elementi lontani hanno interazioni di tolleranza non banali) hanno esponenti critici che dipendono dalla forma specifica dell'interazione, analogamente ai sistemi con interazioni a lungo raggio nella fisica dei materiali.

XI.4 Il crossover tra predicati vaghi e predicati nitidi come transizione di fase

La transizione tra il regime «vago» ($\varepsilon^* > 0$) e il regime «nitido» ($\varepsilon^* = 0$) si manifesta formalmente come una transizione di fase nella struttura del flusso RG. Vicino alla transizione, le quantità fisiche divergono con esponenti caratteristici: la lunghezza di correlazione ξ (la scala su cui due elementi del dominio sono «consapevoli» della struttura globale del predicato) diverge come:

$$\xi \sim |\varepsilon - \varepsilon_c|^{-\nu},$$

dove ε_c è il valore critico della soglia di tolleranza al quale il grafo $G(X, \sim_{\{\varepsilon_c\}})$ transisce dalla connettività alla disconnettività (cioè il valore al quale $\min X$ e $\max X$ si trovano per la prima volta in componenti connesse diverse). Questa è la «transizione di percolazione» del predicato: per $\varepsilon < \varepsilon_c$ il grafo è connesso e il predicato è genuinamente vago; per $\varepsilon > \varepsilon_c$ il grafo può disconnettersi e il predicato può diventare nitido.

La lunghezza di correlazione ξ misura la distanza tipica nel dominio X su cui la vaghezza è «percepibile»: se due elementi $x, y \in X$ hanno $d_G(x, y) \gg \xi$ (dove d_G è la distanza nel grafo di tolleranza), allora i loro assegnamenti ottimali sono praticamente indipendenti. Per ε vicino a ε_c , $\xi \rightarrow \infty$: la vaghezza «permea» l'intero dominio alla scala di correlazione divergente. La funzione di correlazione semantica tra due elementi x, y è:

$$C(x, y) = \langle f^*(x) f^*(y) \rangle - \langle f^*(x) \rangle \cdot \langle f^*(y) \rangle,$$

dove la media $\langle \cdot \rangle$ è rispetto all'ensemble delle strutture di tolleranza compatibili con la scala ε . Si dimostra che $C(x, y) \sim \exp(-d_G(x, y)/\xi)$ per $d_G \gg 1$, con la lunghezza di correlazione ξ che diverge come $|\varepsilon - \varepsilon_c|^{-\nu}$ alla transizione. Il comportamento critico dell'entropia di von Neumann della struttura di vaghezza (già introdotta nella Parte I nell'analogia quantomeccanica) è:

$$S_{vN}(\rho_\varepsilon) \sim \beta \cdot \log |\varepsilon - \varepsilon_c|^{-1} \quad \text{per } \varepsilon \rightarrow \varepsilon_c,$$

dove β è il coefficiente di area della legge logaritmica per l'entropia di entanglement alla criticità (nel campo della teoria delle informazioni quantistiche, questo coefficiente è proporzionale alla carica centrale c della teoria conforme dei campi sottostante). Per la catena di tolleranza unidimensionale alla criticità, $c = 1/2$ (il valore della teoria conforme dei campi di una particella libera massless), coerente con $\nu = 1/2$.

XI.5 Le trasformazioni di dualità e la simmetria tra vago e nitido

Un aspetto delle più originali dell'analogia con il gruppo di rinormalizzazione è la presenza di trasformazioni di dualità che scambiano il regime «vago» con il regime «nitido». Nella fisica delle transizioni di fase, la dualità di Kramers-Wannier scambia l'alta temperatura con la bassa temperatura nel modello di Ising 2D, mappando il regime ordinato (nitido) nel regime disordinato (vago). L'analogo per gli spazi di tolleranza è la trasformazione:

$$D: (X, \sim_\varepsilon) \rightarrow (X, \sim_{\{1-\varepsilon\}}),$$

che scambia la soglia ε con $1 - \varepsilon$. Per $\varepsilon < 1/2$ il grafo $G(X, \sim_\varepsilon)$ è «sparso» (pochi archi, struttura locale); per $\varepsilon > 1/2$ il grafo è «denso» (molti archi, struttura molto connessa). La

dualità scambia i due regimi: un predicato molto vago a soglia ε corrisponde a un predicato «molto nitido» (altamente connesso ma facilmente classificabile) a soglia $1 - \varepsilon$.

Più precisamente: il numero di risoluzione soddisfa $\varepsilon^*(\sim_\varepsilon) = \varepsilon^*(\sim_{1-\varepsilon})$ soltanto se lo spazio è autosimile rispetto a $\varepsilon = 1/2$ — il «punto critico» della dualità. Per la catena lineare con tolleranza uniforme, $\varepsilon^*(\sim_\varepsilon) = \varepsilon/(n-1)$ e $\varepsilon^*(\sim_{1-\varepsilon}) = (1-\varepsilon)/(n-1)$, che sono uguali solo per $\varepsilon = 1/2$. Questo punto critico di dualità è anche il punto di transizione tra il regime in cui il predicato ha una «zona di confine» ben definita ($\varepsilon < 1/2$) e il regime in cui la zona di confine occupa la maggior parte del dominio ($\varepsilon > 1/2$). La «critica di dualità» $\varepsilon = 1/2$ corrisponde semanticamente alla situazione in cui la metà degli archi del grafo di tolleranza completo è inclusa nella struttura di indistinguibilità — il predicato è «half-sharp, half-vague».

Il quadro del gruppo di rinormalizzazione offre infine una prospettiva sui predicati comparativi («più alto di», «più grande di») come predicati al punto fisso instabile del flusso RG. I predicati comparativi hanno struttura di tolleranza invariante per traslazione: la relazione «è più alto di» non dipende dal valore assoluto dell'altezza ma dalla differenza. Questa invarianza è esattamente la condizione che definisce il punto fisso critico del flusso RG — il flusso lascia invariante la struttura invariante per traslazione. I predicati comparativi sono quindi al punto fisso del flusso RG e hanno esponenti critici che dipendono soltanto dalla dimensione dello spazio e dall'invarianza per traslazione, non dai dettagli specifici del predicato. Questa è la ragione matematica per cui i predicati comparativi sembrano «universalmente vaghi» in modo simile tra culture e lingue diverse: la loro struttura di vaghezza è determinata dal punto fisso del flusso RG, che è indipendente dai dettagli culturali e linguistici e dipende soltanto dalla struttura dello spazio percettivo sottostante.

TAVOLA RIASSUNTIVA: I DIZIONARI DI CORRISPONDENZA DELLE PARTI VII–XI

Le cinque sezioni sviluppate in questa sezione aggiuntiva stabiliscono cinque dizionari di corrispondenza precisi tra la teoria degli spazi di tolleranza e altrettanti territori della matematica consolidata. Si riassumono qui le principali corrispondenze, con indicazione del grado di precisione dell'analogia e dei punti in cui essa si interrompe.

Parte VII: Processi stocastici e catene di Markov

Il dizionario è: spazio degli stati $\boxed{\leftrightarrow}$ dominio X del predicato; matrice di transizione P $\boxed{\leftrightarrow}$ struttura di indistinguibilità locale; misura stazionaria $\pi(x) = \deg(x)/(2|E|)$ $\boxed{\leftrightarrow}$ importanza semantica dell'elemento x ; gap spettrale λ_1 $\boxed{\leftrightarrow}$ inverso del costo di coerenza ε^* ; tempo di mescolamento τ_{mix} $\boxed{\leftrightarrow}$ numero di confronti locali necessari per classificare globalmente; funzione armonica h $\boxed{\leftrightarrow}$ funzione di membership ottimale f^* ; energia libera $F(T) = -T \log Z(T)$ $\boxed{\leftrightarrow}$ compromesso vaghezza-nitidezza alla «temperatura semantica» T . L'analogia è precisa per grafi di tolleranza che sono grafi «di Markov reversibili» — cioè per tutti i grafi di tolleranza con la passeggiata casuale semplice. Il punto in cui l'analogia diverge: la catena di Markov è su un dominio statico (X è fisso), mentre nella semantica il

dominio X può cambiare (nuove istanze del predicato possono essere aggiunte). La versione dinamica richiederebbe catene di Markov su spazi degli stati crescenti, un oggetto meno standard.

Parte VIII: Teoria di Morse discreta

Il dizionario è: semplici critici di dimensione 0 $\leftrightarrow$ estremi semantici del predicato (min X e max X); semplici critici di dimensione 1 $\leftrightarrow$ cicli indipendenti del grafo di tolleranza (ostruzioni omologiche); campo vettoriale $V \leftrightarrow$ direzione di crescita ottimale della funzione di membership; cammini di gradiente di $V \leftrightarrow$ cammini monotoni nel grafo di tolleranza; numero di semplici critici $c_k \leftrightarrow$ numeri di Betti β_k ; polinomio di Morse $M(t) \leftrightarrow$ struttura omologica completa del complesso di Sorite. L'analogia è molto precisa per complessi di Sorite 1-dimensionali (grafi). Per complessi di Sorite di dimensione superiore (predicati multidimensionali), la teoria di Morse discreta si applica con più complessità: i semplici critici di dimensione 2 corrispondono a ostruzioni «di secondo ordine» alla coerenza, non previste dalla teoria originale di Sorite. Il punto in cui l'analogia è più ricca della struttura di base: la teoria di Morse discreta fornisce algoritmi computazionali (collasso Morse) per ridurre il complesso di Sorite a una forma minimale, permettendo calcoli omologici molto più efficienti.

Parte IX: Teoria delle rappresentazioni e quiver

Il dizionario è: quiver $A_n \leftrightarrow$ filtrazione di Sorite con $n+1$ soglie; rappresentazione M_k del quiver $\leftrightarrow$ modulo persistente in dimensione k ; indecomponibili $I(i, j) \leftrightarrow$ barre del sorites barcode $[r_i, r_j]$; decomposizione di Krull-Schmidt $\leftrightarrow$ struttura del barcode; teorema di Gabriel $\leftrightarrow$ classificazione completa delle strutture di vaghezza elementari; teorema isometrico $\leftrightarrow$ stabilità del barcode; distanza di interleaving $d_I \leftrightarrow$ distanza di Gromov-Hausdorff tra strutture di tolleranza. L'analogia è la più precisa tra tutte: il sorites barcode è letteralmente la decomposizione di Krull-Schmidt del modulo persistente. Il punto in cui l'analogia rivela qualcosa di nuovo: la dipendenza dal campo $\mathbb{F}$ nella decomposizione produce la nozione di «torsione semantica», un fenomeno genuinamente nuovo che non ha un corrispondente nella teoria classica degli spazi di tolleranza.

Parte X: Geometria simplettica

Il dizionario è: spazio simplettico $(\mathcal{F}_0, \omega) \leftrightarrow$ spazio degli assegnamenti centrati con la forma simplettica discreta; hamiltoniana $H(f) \leftrightarrow$ energia di Dirichlet dell'assegnamento (incoerenza totale); flusso hamiltoniano $\leftrightarrow$ evoluzione degli assegnamenti che conserva l'incoerenza totale; varietà lagrangiana $L_{\text{nitidi}} \leftrightarrow$ spazio degli assegnamenti di predicati nitidi; punto di intersezione lagrangiana $\leftrightarrow$ assegnamento ottimale f^* ; teoria di Floer discreta $\leftrightarrow$ conteggio dei cammini di gradiente tra assegnamenti. Questa analogia è la meno diretta delle cinque: la struttura simplettica è introdotta artificialmente (non emerge da una struttura fisica naturale degli spazi di tolleranza) e il collegamento con l'omologia di Floer richiede ulteriori sviluppi tecnici che esulano dallo scopo di questo volume. Il punto in cui l'analogia è meno precisa: la geometria simplettica richiede una struttura differenziabile o almeno una struttura lineare «liscia», mentre gli spazi di tolleranza finiti sono oggetti discreti e combinatoriali.

Parte XI: Gruppo di rinormalizzazione

Il dizionario è: decimazione D_b $\leftrightarrow$ aggregazione di elementi vicini nel dominio; flusso RG $\leftrightarrow$ evoluzione della struttura di vaghezza al variare della scala di osservazione; punto fisso stabile $\leftrightarrow$ predicato nitido (attrattore); punto fisso instabile $\leftrightarrow$ predicato con struttura autosimile (catena infinita uniforme); esponente critico ν $\leftrightarrow$ velocità di divergenza del numero di risoluzione alla transizione; lunghezza di correlazione ξ $\leftrightarrow$ scala su cui la vaghezza è percepibile nel dominio; dualità di Kramers-Wannier $\leftrightarrow$ scambio tra regime vago e regime nitido; punto fisso della dualità $\varepsilon = 1/2$ $\leftrightarrow$ predicato «mezzo-vago»; carica centrale $c = 1/2$ $\leftrightarrow$ esponente critico $\nu = 1/2$ per predicati unidimensionali. L'analogia con il RG è la più programmatica delle cinque: fornisce un quadro qualitativo ricco ma i calcoli espliciti degli esponenti critici e della struttura delle classi di universalità richiedono sviluppi tecnici aggiuntivi, in particolare la costruzione della teoria conforme dei campi discreta associata al complesso di Sorite. Il punto in cui l'analogia apre le strade più interessanti: la nozione di universalità della vaghezza — che predicati con strutture microscopiche molto diverse hanno lo stesso esponente critico ν — offre una spiegazione matematica del fatto empiricamente osservato che la vaghezza dei predicati gradualmente sembra «universale» tra culture diverse, indipendentemente dai dettagli lessicali e grammaticali di ciascuna lingua.

L'insieme delle cinque analogie stabilisce che la teoria degli spazi di tolleranza si inserisce naturalmente in cinque grandi filoni della matematica contemporanea, ciascuno dei quali contribuisce strumenti e risultati trasferibili alla semantica formale. Nessuna delle singole analogie è completa: ciascuna cattura aspetti diversi della struttura di vaghezza e identifica limiti distinti. Il quadro complessivo — che emerge dalla convergenza delle cinque prospettive — è quello di una teoria della vaghezza linguistica radicata in strutture matematiche profonde e già ampiamente sviluppate, con la conseguenza che i progressi in ciascuno dei cinque territori si traducono automaticamente in progressi nella comprensione della struttura formale dei predicati vaghi.

RIFERIMENTI AGGIUNTIVI PER LA SEZIONE SPECIALE

Meccanica quantistica e quantum cognition: R. Bhatia, *Matrix Analysis*, Springer (1997) [per la distanza di Bures e la fidelity]. J. Busemeyer e P. Bruza, *Quantum Models of Cognition and Decision*, Cambridge University Press (2012). D. Aerts, S. Sozzo e T. Veloz, "New Axiomatics for the Formation of Quantum Structures in Cognition", *Topoi* 34 (2015). J. S. Bell, "On the Einstein-Podolsky-Rosen Paradox", *Physics* 1(3), 195-200 (1964). J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, "Proposed Experiment to Test Local Hidden-Variable Theories", *Physical Review Letters* 23, 880-884 (1969). B. Tsirelson, "Quantum Generalizations of Bell's Inequality", *Letters in Mathematical Physics* 4, 93-100 (1980). S. Kochen e E. P. Specker, "The Problem of Hidden Variables in Quantum Mechanics", *Journal of Mathematics and Mechanics* 17, 59-87 (1967). G. Birkhoff e J. von Neumann, "The Logic of Quantum Mechanics", *Annals of Mathematics* 37, 823-843 (1936).

Fisica statistica e transizioni di fase: R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press (1982). L. Onsager, "Crystal Statistics I. A Two-Dimensional

Model with an Order-Disorder Transition", *Physical Review* 65, 117-149 (1944). P. C. Hohenberg e B. I. Halperin, "Theory of Dynamic Critical Phenomena", *Reviews of Modern Physics* 49, 435-479 (1977). D. Stauffer e A. Aharony, *Introduction to Percolation Theory*, Taylor & Francis (1992). K. G. Wilson e J. Kogut, "The Renormalization Group and the ϵ Expansion", *Physics Reports* 12(2), 75-199 (1974). V. L. Berezinskii, "Destruction of Long-Range Order in One-Dimensional and Two-Dimensional Systems Having a Continuous Symmetry Group", *Soviet Physics JETP* 32, 493-500 (1971). J. M. Kosterlitz e D. J. Thouless, "Ordering, Metastability and Phase Transitions in Two-Dimensional Systems", *Journal of Physics C: Solid State Physics* 6, 1181-1203 (1973).

Teoria dell'informazione: C. E. Shannon, "A Mathematical Theory of Communication", *Bell System Technical Journal* 27, 379-423 (1948). C. E. Shannon, "The Zero Error Capacity of a Noisy Channel", *IRE Transactions on Information Theory* 2(3), 8-19 (1956). C. E. Shannon, "Coding Theorems for a Discrete Source with a Fidelity Criterion", *IRE National Convention Record Part 4*, 142-163 (1959). L. Lovász, "On the Shannon Capacity of a Graph", *IEEE Transactions on Information Theory* 25(1), 1-7 (1979). R. W. Hamming, "Error Detecting and Error Correcting Codes", *Bell System Technical Journal* 29, 147-160 (1950). I. Csiszár e J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Cambridge University Press (2011).

Geometria riemanniana discreta: Y. Ollivier, "Ricci Curvature of Markov Chains on Metric Spaces", *Journal of Functional Analysis* 256(3), 810-864 (2009). Y. Lin, L. Lu, S. T. Yau, "Ricci Curvature of Graphs", *Tohoku Mathematical Journal* 63(4), 605-627 (2011). C. Villani, *Optimal Transport: Old and New*, Springer (2009). L.-H. Lim e R. Orsi, "The Curvature of a Graph", *Journal of Pure and Applied Algebra* (2023). D. Bakry e M. Émery, "Diffusions Hypercontractives", in *Séminaire de Probabilités XIX, Lecture Notes in Mathematics* 1123, 177-206, Springer (1985). M. Gromov, *Metric Structures for Riemannian and Non-Riemannian Spaces*, Birkhäuser (1999). P. Delmotte, "Parabolic Harnack Inequality and Estimates of Markov Chains on Graphs", *Revista Matemática Iberoamericana* 15, 181-232 (1999).

Teoria delle categorie e topos: S. MacLane e I. Moerdijk, *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*, Springer (1992). P. T. Johnstone, *Sketches of an Elephant: A Topos Theory Compendium*, Oxford University Press (2002). F. W. Lawvere e M. Tierney, "Quantifiers and Sheaves", in *Actes du Congrès International des Mathématiciens*, 329-334 (1970). G. Grätzer, *General Lattice Theory*, Birkhäuser (1978). J.-L. Bell, *Toposes and Local Set Theories: An Introduction*, Clarendon Press (1988).

Analisi armonica sui grafi: R. R. Coifman e M. Maggioni, "Diffusion Wavelets", *Applied and Computational Harmonic Analysis* 21, 53-94 (2006). D. K. Hammond, P. Vandergheynst e R. Gribonval, "Wavelets on Graphs via Spectral Graph Theory", *Applied and Computational Harmonic Analysis* 30, 129-150 (2011). P. Li e S. T. Yau, "On the Parabolic Kernel of the Schrödinger Operator", *Acta Mathematica* 156, 153-201 (1986). F. Bauer, P. Horn, Y. Lin, G. Lippner, D. Mangoubi e S. T. Yau, "Li-Yau Inequality on Graphs", *Journal of Differential Geometry* 99, 359-405 (2015).

Teoria dei giochi e struttura combinatoria: J. Hintikka, "Modality as Referential Multiplicity", *Ajatus* 20, 49-64 (1957). L. Lorenzen, "Logik und Agon", *Atti del XII Congresso Internazionale di Filosofia*, 187-194 (1961). M. Aigner e M. Fromme, "A Game of Cops and Robbers on Graphs", *Discrete Applied Mathematics* 8(1), 1-12 (1984). H. L. Bodlaender, "A Partial k -Arboretum of Graphs with Bounded Treewidth", *Theoretical Computer Science* 209(1-2), 1-45 (1998).

Termodinamica dell'informazione: E. T. Jaynes, "Information Theory and Statistical Mechanics", Physical Review 106, 620-630 (1957). R. Landauer, "Irreversibility and Heat Generation in the Computing Process", IBM Journal of Research and Development 5, 183-191 (1961). C. H. Bennett, "The Thermodynamics of Computation: A Review", International Journal of Theoretical Physics 21, 905-940 (1982).

PARTE XII. GEOMETRIA DELL'INFORMAZIONE E METRICA DI FISHER SULLO SPAZIO DEI PREDICATI VAGHI

XII.1 Lo spazio statistico dei predicati vaghi e la metrica di Fisher

La geometria dell'informazione, fondata da Rao (1945) e sistematizzata da Amari e Nagaoka nel loro testo fondamentale (2000), studia le proprietà differenziali degli spazi di distribuzioni di probabilità, dotandoli di una struttura riemanniana naturale. La metrica riemanniana associata — la metrica di Fisher — misura la «distinguibilità» tra distribuzioni di probabilità vicine: due distribuzioni sono molto vicine se e solo se l'esperimento statistico più favorevole non riesce a distinguerle con affidabilità migliore di quella del caso. Questo approccio si trasferisce in modo naturale agli spazi di predicati vaghi, dove la «distinguibilità» tra predicati sostituisce la distinguibilità tra distribuzioni.

Si considera lo spazio dei predicati vaghi su un dominio fisso $X = \{x_1 < x_2 < \dots < x_n\}$ con tolleranza $\sim$ data. Un predicato è rappresentato da una funzione di membership $f: X \rightarrow [0,1]$ ε^* -coerente, monotona, con $f(x_1) = 0$ e $f(x_n) = 1$. Lo spazio di tali funzioni è il poliedro convesso $\mathcal{F}_{\varepsilon^*} \subset [0,1]^n$ descritto nel Capitolo 10. Per dotarlo di una struttura di spazio statistico, si interpreta ogni funzione f come il parametro di una famiglia di distribuzioni di Bernoulli: per ogni elemento $x \in X$, la variabile aleatoria $B(x) \sim \text{Bernoulli}(f(x))$ modella il fatto che un parlante scelga a caso, in modo indipendente per ogni x , se classificare x come «vero» (con probabilità $f(x)$) o «falso» (con probabilità $1 - f(x)$).

Formalmente, si considera la famiglia di distribuzioni di probabilità su $\{0,1\}^n$ parametrizzata da $f \in \mathcal{F}_{\varepsilon^*}$:

$$P_{\mathbf{f}}(\sigma_1, \sigma_2, \dots, \sigma_n) = \prod_i f(x_i)^{\sigma_i} \cdot (1 - f(x_i))^{1-\sigma_i}, \quad \sigma_i \in \{0, 1\}$$

dove la distribuzione fattorizza sulla dipendenza indipendente per ciascun elemento (questo è il modello più semplice; si analizzeranno estensioni con dipendenze tra elementi vicini nella sezione XII.4). La matrice di informazione di Fisher del modello è la matrice $n \times n$:

$$I(f)_{ij} = E_f[\partial/\partial f_i \log P_f \cdot \partial/\partial f_j \log P_f]$$

dove $f_i = f(x_i)$. Per il modello fattorizzato di Bernoulli, il calcolo è diretto: poiché $\log Pf = \sum_i [\sigma_i \log f_i + (1-\sigma_i) \log(1-f_i)]$, la derivata parziale rispetto a f_i è $\partial/\partial f_i \log Pf = \sigma_i/f_i - (1-\sigma_i)/(1-f_i)$. I termini misti $i \neq j$ non contribuiscono all'aspettativa per l'indipendenza, e il termine diagonale è:

$$I(f)_{ii} = E_f[(\sigma_i/f_i - (1-\sigma_i)/(1-f_i))^2] = 1/(f_i(1-f_i))$$

Dunque la matrice di informazione di Fisher del modello fattorizzato di Bernoulli sui predicatori è la matrice diagonale:

$$I(f) = \text{diag}(1/(f_1(1-f_1)), 1/(f_2(1-f_2)), \dots, 1/(f_n(1-f_n)))$$

La metrica riemanniana di Fisher sullo spazio $\mathcal{F}_{\varepsilon^*}$ è la forma quadratica:

$$g_f(\delta f, \delta f') = \sum_i \delta f_i \delta f'_i / (f_i(1-f_i))$$

Questa è la metrica naturale sul $(n-1)$ -simpleso di probabilità, ristretta al poliedro convesso $\mathcal{F}_{\varepsilon^*}$. La geometria della metrica di Fisher sulla varietà delle distribuzioni di Bernoulli è quella della sfera unitaria (sotto la trasformazione $\arcsin(\sqrt{f})$, la metrica di Fisher diventa la metrica euclidea standard): questo è il fondamento della «trasformazione di arcsin» della statistica. Per il poliedro $\mathcal{F}_{\varepsilon^*}$, la struttura non è semplicemente quella sferica perché i vincoli del poliedro (monotonia, ε^* -coerenza) introducono una struttura di varietà con bordo.

Definizione XII.1 (Geodetiche dello spazio dei predicatori vaghi). Una geodetica $\gamma: [0,1] \rightarrow \mathcal{F}_{\varepsilon^*}$ rispetto alla metrica di Fisher g è una curva di lunghezza minima che connette due predicatori $f_0, f_1 \in \mathcal{F}_{\varepsilon^*}$. La lunghezza di Fisher della curva γ è: $L_F(\gamma) = \int_0^1 \sqrt{(g_{\gamma(t)}(\gamma'(t), \gamma'(t)))} dt = \int_0^1 \sqrt{(\sum_i (\gamma'_i(t))^2 / (\gamma_i(t)(1-\gamma_i(t))))} dt$

Sotto la trasformazione $\varphi_i = 2 \arcsin(\sqrt{f_i}) \in [0, \pi]$ (la trasformazione di arcsin che «linearizza» la metrica di Fisher per il Bernoulli), la lunghezza di Fisher diventa la distanza euclidea nello spazio trasformato:

$$L_F(\gamma) = \int_0^1 \|\varphi'(\gamma(t))\|_2 dt, \quad \text{dove } \varphi(f)_i = 2 \arcsin(\sqrt{f_i})$$

La distanza di Fisher tra due predicatori f_0 e f_1 in $\mathcal{F}_{\varepsilon^*}$ è:

$$d_F(f_0, f_1) = \|2 \arcsin(\sqrt{f_0}) - 2 \arcsin(\sqrt{f_1})\|_2 = 2 \sqrt{(\sum_i (\arcsin(\sqrt{f_{0i}}) - \arcsin(\sqrt{f_{1i}}))^2)}$$

per geodetiche interne al poliedro; se la geodetica euclidea nello spazio φ non è contenuta in $\varphi(\mathcal{F}_{\varepsilon^*})$, la distanza si calcola con la proiezione geodetica sul bordo del poliedro trasformato.

XII.2 Il bound di Cramér-Rao per la stima della vaghezza

Il teorema di Cramér-Rao (Cramér, 1946; Rao, 1945) è uno dei risultati fondamentali della statistica matematica: stabilisce un limite inferiore alla varianza di qualunque stimatore non distorto di un parametro. Applicato allo spazio dei predicati vaghi, fornisce un limite inferiore alla precisione con cui la funzione di membership di un predicato può essere stimata da dati empirici (classificazioni di istanze da parte di parlanti).

Si supponga di osservare m classificazioni indipendenti di un'istanza x_i : m parlanti classificano indipendentemente x_i come «vero» (con probabilità $f(x_i)$) o «falso». Lo stimatore naturale è la frequenza empirica $\hat{f}_i = (\text{numero di «vero»})/m$. Per il modello di Bernoulli, la varianza dello stimatore è:

$$\text{Var}(\hat{f}_i) = f(x_i)(1-f(x_i))/m$$

Il bound di Cramér-Rao stabilisce che per ogni stimatore non distorto T di $f(x_i)$ da m osservazioni:

$$\text{Var}(T) \geq 1/(m \cdot I(f)_{ii}) = f(x_i)(1-f(x_i))/m$$

La frequenza empirica realizza il bound: è lo stimatore di massima verosimiglianza ed è efficiente. Per stimare simultaneamente l'intera funzione di membership f su tutto X , si utilizza la matrice di informazione di Fisher $I(f)$ e il bound matriciale di Cramér-Rao:

$$\text{Cov}(\hat{T}) \geq (1/m) \cdot I(f)^{-1} = (1/m) \cdot \text{diag}(f_1(1-f_1), \dots, f_n(1-f_n))$$

nel senso che $\text{Cov}(\hat{T}) - I(f)^{-1}/m$ è una matrice semidefinita positiva. Questo risultato ha un'interpretazione semantica: le istanze x_i con $f(x_i)$ vicino a $1/2$ (la zona di confine del predicato) hanno $I(f)_{ii} = 1/(f(x_i)(1-f(x_i)))$ vicino al suo valore minimo di 4, e richiedono il maggior numero di classificazioni per essere stimate con precisione data; le istanze con $f(x_i)$ vicino a 0 o 1 hanno $I(f)_{ii} \rightarrow \infty$ (la zona chiara del predicato) e sono stimate con grande precisione anche con poche osservazioni.

Teorema XII.1 (Precisione minima nella stima della vaghezza). Sia $f^* \in \mathcal{F}_{\{\varepsilon^*\}}$ il predicato ottimale (Capitolo 5). Il numero minimo di classificazioni m_i necessario per stimare $f^*(x_i)$ con errore quadratico medio $\leq \delta^2$ è: $m_i \geq f^*(x_i)(1-f^*(x_i))/\delta^2 \leq 1/(4\delta^2)$. Il numero totale di classificazioni necessario per stimare f^* su tutto X con errore massimo δ è: $m_{\text{tot}} \geq \sum_i f^*(x_i)(1-f^*(x_i))/\delta^2 \leq n/(4\delta^2)$. La zona di confine $B(P) = \{x : |f^*(x) - 1/2| < \eta\}$ contribuisce la maggior parte delle classificazioni necessarie: $m_{\text{confine}}/m_{\text{tot}} \rightarrow 1$ per $\eta \rightarrow 1/2$.

Dimostrazione. La prima disuguaglianza segue direttamente dal bound di Cramér-Rao per lo stimatore di $f^*(x_i)$. La seconda dalla somma su tutti i vertici. Per la terza parte, si osserva che $f^*(x_i)(1-f^*(x_i))$ assume il massimo $1/4$ per $f^*(x_i) = 1/2$ e decresce quadraticamente allontanandosi da $1/2$. Per f^* lineare (predicato di Łukasiewicz), $f^*(x_i) = (i-1)/(n-1)$, e il contributo dell'elemento x_i alla somma totale è $(i-1)/(n-1) \cdot (n-i)/(n-1) = (i-1)(n-i)/(n-1)^2$. La somma sull'intera zona di confine $|f^*(x_i) - 1/2| < \eta$ domina la somma totale per $\eta \rightarrow 1/2$. ■

XII.3 La curvatura di Fisher e la struttura della zona di confine

La curvatura riemanniana della varietà statistica $(\mathcal{F}_{\{\varepsilon^*\}}, g_F)$ misura quanto la geometria dello spazio dei predicati si discosta dalla piattezza euclidea. Per la metrica di Fisher del modello di Bernoulli (senza i vincoli del poliedro), la curvatura scalare è:

$$R_F = \sum_i R_{iiii} = \sum_i \left(\frac{1}{2} \right) \cdot \left(\frac{\partial^2}{\partial f_i^2} \log f(x_i) (1-f(x_i)) \right)$$

Nel modello fattorizzato, la curvatura sezionale per ogni coppia (i, j) con $i \neq j$ è zero (il modello è un prodotto di manifold 0-dimensional). La curvatura compare invece lungo la direzione del bordo del poliedro: la condizione di ε^* -coerenza $|f_i - f_j| \leq \varepsilon^*$ per coppie tolleranti introduce un «wall» nel poliedro, e la curvatura del bordo è determinata dal curvature tensor della metrica di Fisher ristretta al bordo. Per un vincolo lineare della forma $f_j - f_i = \varepsilon^*$ (il bordo del poliedro in direzione ij), la curvatura geodetica del bordo rispetto alla metrica di Fisher è:

$$\kappa_{\text{geodesica}} = \left(\frac{1}{2(f_i(1-f_i))^{3/2}} \right) \cdot (1 - 2f_i) \cdot \frac{1}{\sqrt{f_j(1-f_j)}}$$

che è negativa (concava verso l'interno del poliedro) per $f_i > 1/2$ e positiva (convessa) per $f_i < 1/2$. Il punto critico dove la curvatura cambia segno è esattamente $f_i = 1/2$: il punto mediano della funzione di membership, che corrisponde al confine semantico del predicato. Questa analisi differenziale recupera, mediante la geometria di Fisher, la localizzazione della zona di confine già ottenuta geometricamente (Capitolo 18) tramite la curvatura di Ricci di Ollivier.

Definizione XII.2 (Curvatura semantica di Fisher). Per ogni elemento $x_i \in X$, la curvatura semantica di Fisher è: $\kappa_F(x_i) = 1 - 2f^*(x_i) / (f^*(x_i)(1-f^*(x_i)))$. La curvatura semantica è positiva nella regione chiaramente falsa ($f^*(x_i) < 1/2$), negativa nella zona di confine ($f^*(x_i) \approx 1/2$), e diverge a $\pm\infty$ per $f^*(x_i) \rightarrow 0$ o 1 . Il punto di inflessione $\kappa_F = 0$ è raggiunto per $f^*(x_i) = 1/2$ e identifica il centro della zona di confine.

XII.4 Modelli con dipendenze: la metrica di Fisher per il modello di Ising semantico

Il modello fattorizzato di Bernoulli tratta le classificazioni di elementi diversi come indipendenti. Una modellizzazione più realistica tiene conto delle dipendenze tra classificazioni di elementi vicini: se un parlante classifica x_i come «vero», è più probabile che classifichi anche x_{i+1} come «vero» (dipendenza locale). Il modello naturale per tale dipendenza è il modello di Gibbs (o modello di Ising generalizzato) sulla struttura del grafo di tolleranza.

Si definisce il modello di Ising semantico come la distribuzione di probabilità sul cubo $\{0,1\}^n$:

$$P_{\{f, \beta\}}(\sigma) = Z(f, \beta)^{-1} \exp(\sum_i h_i \sigma_i + \beta \sum_{\{(i, j) \in E\}} J_{ij} \sigma_i \sigma_j)$$

dove $h_i = \log(f(x_i)/(1-f(x_i)))$ è il campo magnetico locale corrispondente alla funzione di membership $f(x_i)$, $\beta > 0$ è la «temperatura inversa» semantica (β grande corrisponde a classificazioni fortemente coerenti), $J_{ij} = \varepsilon^*(\sim)/(|i-j|+1)$ è la costante di accoppiamento tra gli elementi x_i e x_j (decrescente con la distanza ordinale), e $Z(f, \beta) = \sum_{\sigma} \exp(\cdot)$ è la funzione di partizione. Per $\beta \rightarrow 0$ (temperatura alta), il modello tende al modello di Bernoulli indipendente. Per $\beta \rightarrow \infty$ (temperatura zero), il modello tende a essere concentrato sulle configurazioni che

massimizzano l'energia — cioè sulle configurazioni «coerenti» nel senso che elementi adiacenti nel grafo di tolleranza ricevono lo stesso valore.

La matrice di informazione di Fisher del modello di Ising semantico è la somma di un termine diagonale (contributo dei campi locali h_i) e di un termine di interazione (contributo degli accoppiamenti J_{ij}):

$$I_{\{f, \beta\}}(f, \beta)_{ij} = \text{Cov}_{\{f, \beta\}}(\partial/\partial h_i \log P_{\{f, \beta\}}, \partial/\partial h_j \log P_{\{f, \beta\}}) \\ = \text{Cov}_{\{f, \beta\}}(\sigma_i, \sigma_j)$$

dove $\text{Cov}_{\{f, \beta\}}(\sigma_i, \sigma_j) = \langle \sigma_i \sigma_j \rangle_{\{f, \beta\}} - \langle \sigma_i \rangle_{\{f, \beta\}} \langle \sigma_j \rangle_{\{f, \beta\}}$ è la funzione di correlazione connessa del modello di Ising semantico. Per β piccolo (alta temperatura), la funzione di correlazione decade esponenzialmente con la distanza: $\text{Cov}(\sigma_i, \sigma_j) \sim \exp(-|i-j|/\xi_F)$ dove ξ_F è la lunghezza di correlazione di Fisher. La relazione tra ξ_F e il numero di risoluzione ε^* è:

$$\xi_F \approx 1/(2\beta \cdot \varepsilon^*(\sim)) \quad \text{per } \beta \cdot \varepsilon^*(\sim) \ll 1 \text{ (regime ad alta temperatura)}$$

Questa relazione stabilisce che la lunghezza di correlazione di Fisher (la scala entro cui le classificazioni di parlanti diversi sono correlate) è inversamente proporzionale al prodotto della temperatura inversa semantica β e del numero di risoluzione ε^* . Predicati più vaghi (ε^* piccolo) hanno lunghezze di correlazione più lunghe — le classificazioni di parlanti diversi sono più correlate, nel senso che la struttura del predicato si propaga su distanze più lunghe nel dominio.

XII.5 La divergenza di Kullback-Leibler tra predicati e la topologia dello spazio informativo

La divergenza di Kullback-Leibler (Kullback e Leibler, 1951) tra due distribuzioni P_f e P_g rappresenta la «perdita di informazione» che si subisce usando g invece di f come modello:

$$KL(P_f \parallel P_g) = \sum_i [f_i \log(f_i/g_i) + (1-f_i) \log((1-f_i)/(1-g_i))]$$

La divergenza KL non è simmetrica, ma la sua versione simmetrizzata (la divergenza di Jeffrey) $KL_{\text{sym}}(f, g) = (KL(f \parallel g) + KL(g \parallel f))/2$ e la radice quadrata della metrica χ^2 (la metrica di Pearson) sono entrambe strettamente collegate alla metrica di Fisher: per f e g vicini, $KL(P_f \parallel P_g) \approx (1/2) gF(f-g, f-g)$, cioè la divergenza KL approssima la distanza quadratica di Fisher al secondo ordine. La topologia dello spazio $(\mathcal{F}_{\varepsilon^*}, KL)$ è pertanto quella di una varietà riemanniana per distanze piccole e di uno spazio metrico generale per distanze grandi.

La struttura globale della topologia KL ha implicazioni importanti per il confronto di predicati vaghi. Due predicati f e g sono «informativamente equivalenti» se $KL(P_f \parallel P_g) = KL(P_g \parallel P_f) = 0$, il che (per il modello di Bernoulli) implica $f = g$: la divergenza KL classifica i predicati in modo più fine della distanza bottleneck del sorites barcode (che distingue solo i predicati con strutture omologiche diverse). D'altra parte, la distanza KL è più fine di quella bottleneck ma meno fine della funzione di membership stessa: due predicati con KL piccola hanno funzioni di membership simili, ma due predicati con la stessa funzione di membership sono identici.

Si definisce la *palla di informazione* di raggio r attorno a un predicato f :

$$B_{KL}(f, r) = \{g \in \mathcal{F}_{\{\varepsilon^*\}} : KL(P_f \parallel P_g) \leq r\}$$

Il volume di questa palla rispetto alla misura di Liouville della metrica di Fisher è asintoticamente (per r piccolo):

$$\text{Vol}(B_{KL}(f, r)) \approx (\pi r)^{n/2} / \Gamma(n/2 + 1) \cdot (\det I(f))^{-1/2}$$

dove $\det I(f) = \prod_i 1/(f_i(1-f_i))$ è il determinante della matrice di Fisher. Il volume è massimo nella zona di confine dove $f_i \approx 1/2$ e $\det I(f) \approx 4^n$ è minimo, e minimo agli estremi dove $f_i \approx 0$ o 1 e $\det I(f) \rightarrow \infty$. Questo riflette l'intuizione che nella zona di confine i predicati sono «densi» — vicini l'uno all'altro nello spazio informativo — mentre agli estremi i predicati netti sono «sparsi» e facilmente distinguibili.

Teorema XII.2 (Entropia di Fisher della zona di confine). Sia $B(P) = \{x_i : |f^*(x_i) - 1/2| < \eta\}$ la zona di confine del predicato ottimale f^* . L'entropia di Fisher dell'assegnamento ottimale f^* , definita come: $S_F(f^*) = -\int_{\mathcal{F}_{\{\varepsilon^*\}}} \log p(f^*; g) \cdot p(f^*; g) d\text{Vol}_F(g)$ dove $p(f^*; g)$ è la densità gaussiana di Fisher centrata in f^* , è: $S_F(f^*) = n/2 (1 + \log(2\pi)) - (1/2) \log \det I(f^*) = n/2 (1 + \log(2\pi)) + (1/2) \sum_i \log(f_i^*(1-f_i^*))$. L'entropia è massima quando tutti gli $f_i^* = 1/2$ (vaghezza massima) e minima agli estremi (predicato netto).

PARTE XIII. LA METRICA DI WASSERSTEIN E IL TRASPORTO OTTIMALE TRA STRUTTURE DI VAGHEZZA

XIII.1 Formulazione del problema di trasporto ottimale per i predicati vaghi

La teoria del trasporto ottimale, sviluppata originariamente da Monge (1781) nel contesto del problema di trasportare terra da un sito a un altro al minimo costo, e riformulata nella versione rilassata da Kantorovich (1942), ha avuto negli ultimi decenni uno sviluppo straordinario con applicazioni in geometria differenziale, teoria della probabilità, equazioni alle derivate parziali e analisi dei dati. Villani (2003, 2009) ha sistematizzato la teoria in una forma che la rende applicabile a spazi metrici generali. Il contributo di questo capitolo è applicare il trasporto ottimale allo spazio dei predicati vaghi, ottenendo la metrica di Wasserstein come strumento per misurare la «distanza semantica» tra predicati in modo più sensibile alla geometria del dominio rispetto alla distanza bottleneck del sorites barcode.

Si fissa un dominio $X = \{x_1 < \dots < x_n\}$ con struttura di tolleranza $\sim$. Due predicati vaghi $f, g: X \rightarrow [0,1]$ sono interpretati come distribuzioni di massa: $f_i = f(x_i)$ è la massa posiziona in x_i (la

probabilità che x_i sia classificato come «vero» secondo il predicato f). Per normalizzare, si considerano le funzioni di membership come densità rispetto alla misura uniforme su X : la misura $\mu_f = (1/n) \sum_i f(x_i) \delta_{\{x_i\}}$ è la distribuzione di probabilità associata a f sul dominio X , con massa totale $\mu_f(X) = (1/n) \sum_i f(x_i) = \bar{f}$ (la media della funzione di membership).

Poiché due predicati f e g in generale hanno masse totali diverse ($\bar{f} \neq \bar{g}$), non si può applicare direttamente il trasporto ottimale (che richiede che le distribuzioni sorgente e destinazione abbiano la stessa massa totale). Si considera invece il problema di trasporto tra la distribuzione differenziale: si definisce la misura con segno $\nu\{f, g\} = \mu_f - \mu_g$ e si decompone $\nu\{f, g\} = \nu^+ - \nu^-$ nella sua parte positiva e negativa. Il problema di trasporto ottimale unbiased tra f e g è il problema di minimizzare il costo di trasporto da ν^+ a ν^- con funzione di costo $c(x_i, x_j) = d_G(x_i, x_j)$ (distanza nel grafo di tolleranza).

Per predicati f, g con la stessa media ($\bar{f} = \bar{g} = 1/2$, corrispondenti a predicati «bilanciati»), la costruzione si semplifica: la metrica di Wasserstein-1 tra f e g è:

$$W_1(f, g) = \min_{\pi \in \Pi(\mu_f, \mu_g)} \sum_{i,j} c(x_i, x_j) \cdot \pi(x_i, x_j)$$

dove $\Pi(\mu_f, \mu_g)$ è l'insieme dei piani di trasporto — le distribuzioni di probabilità π su $X \times X$ con marginali μ_f e μ_g rispettivamente. Per la metrica di costo $c = d_G$ (distanza nel grafo), la metrica W_1 si calcola come il valore ottimale del programma lineare:

$$W_1(f, g) = \max_{\varphi: X \rightarrow \mathbb{R}, |\varphi(x_i) - \varphi(x_j)| \leq d_G(x_i, x_j)} \sum_i \varphi(x_i) (f_i - g_i) / n$$

per la dualità di Kantorovich. La funzione φ nella formulazione duale è una «funzione di potenziale di trasporto» che soddisfa la condizione di 1-Lipschitz rispetto alla metrica del grafo.

XIII.2 La metrica di Wasserstein-2 e la struttura della geodetiche ottimali

La metrica di Wasserstein-2 è la versione quadratica del trasporto ottimale, con funzione di costo $c(x_i, x_j) = d_G(x_i, x_j)^2$:

$$W_2(f, g)^2 = \min_{\pi \in \Pi(\mu_f, \mu_g)} \sum_{i,j} d_G(x_i, x_j)^2 \cdot \pi(x_i, x_j)$$

Lo spazio $(\mathcal{F}\{\varepsilon^*\}, W_2)$ è uno spazio metrico (o quasi-metrico nel caso non bilanciato) con struttura riemanniana formale analoga a quella dello spazio di Wasserstein $P_2(M)$ su una varietà M , sviluppata da Otto (2001) e Villani (2003). La struttura di spazio metrico di $(\mathcal{F}\{\varepsilon^*\}, W_2)$ permette di definire le *geodetiche di McCann*: una geodetica da f a g è la famiglia di predicati interpolati $f_t = ((1-t)\varphi_f + t\varphi_g) \# \mu$ dove φ_f, φ_g sono le funzioni di trasporto ottimale (mappe di Brenier nel caso di spazi euclidei, convex combinations nel caso discreto). Nel caso discreto, la geodetica ottimale trasporta la massa da μ_f a μ_g seguendo il piano di trasporto ottimale π^* :

$$f_t(x_k) = \sum_i \pi^*(x_i, x_j) \cdot \mathbb{1}[(1-t)x_i + tx_j = x_k] / \mu_{\text{uniform}}(x_k)$$

dove l'interpolazione lineare $(1-t)x_i + tx_j$ è da interpretarsi nel senso del grafo: si percorre il cammino minimo da x_i a x_j in $G(X, \sim)$ arrestandosi alla frazione t del percorso. Per la catena lineare con metrica ordinale $d(x_i, x_j) = |i-j|$, la geodetica di McCann è particolarmente esplicita: la massa in

x_i si muove verso la massa in x_j con velocità costante, e al tempo t si trova nell'elemento $x_{\lfloor i + t(j-i) \rfloor}$ (arrotondato all'intero più vicino).

Teorema XIII.1 (Contrazione della metrica di Wasserstein sotto il flusso di Sorite). Sia $F_t: \mathcal{F}_{\varepsilon^*} \rightarrow \mathcal{F}_{\varepsilon^*}$ il flusso dell'equazione del calore semantica (introdotta nella Parte XIV), che manda ogni predicato f nel predicato 'smussato' $f_t = e^{\{t \cdot \Delta_{\sim}\}} f$, dove $\Delta_{\sim}$ è il laplaciano del grafo di tolleranza. Allora il flusso F_t è una contrazione rispetto alla metrica W_2 : $W_2(F_t f, F_t g) \leq e^{\{-\lambda_1 t\}} \cdot W_2(f, g)$ per ogni $t \geq 0$, dove $\lambda_1 > 0$ è il gap spettrale del laplaciano $L = D - A$ del grafo di tolleranza. Il tasso di contrazione $e^{\{-\lambda_1 t\}}$ coincide con il tasso di decadimento dell'energia di Dirichlet calcolato nel Capitolo 8.

Dimostrazione (sketch). Il flusso del calore $e^{\{t\Delta\}}$ sui grafi soddisfa la disuguaglianza di Bakry-Émery $CD(K, \infty)$ con $K = \lambda_1/2$ (la costante di Ricci discreta di Ollivier del grafo). Per spazi che soddisfano la condizione $CD(K, \infty)$, la teoria sviluppata da Erbar e Maas (2012) e da Mielke (2011) nell'ambito della struttura metrica del trasporto ottimale garantisce la contrazione della metrica W_2 con tasso $e^{\{-Kt\}}$: $W_2(e^{\{t\Delta\}} f, e^{\{t\Delta\}} g) \leq e^{\{-(\lambda_1/2)t\}} W_2(f, g)$. Il fattore $1/2$ nel tasso di contrazione rispetto a λ_1 emerge dalla relazione $K = \lambda_1/2$ tra il bound di curvatura di Bakry-Émery e il gap spettrale. La stima più precisa $W_2(F_t f, F_t g) \leq e^{\{-\lambda_1 t\}} W_2(f, g)$ si ottiene dal calcolo diretto per grafi: la norma L_2 della differenza $f_t - g_t$ decade come $e^{\{-\lambda_1 t\}}$ $\|f - g\|_2$, e per grafo bipartito la distanza W_2 è dominata dalla norma L_2 a meno di un fattore costante dipendente dal diametro. ■

XIII.3 Il teorema di De Giorgi-Rudin-Osher per predicati vaghi e la variazione totale semantica

La variazione totale di una funzione $f: X \rightarrow \mathbb{R}$ su un grafo $G = (V, E)$ è:

$$TV_G(f) = \sum_{\{i,j\} \in E} w(i,j) \cdot |f(x_i) - f(x_j)|$$

dove $w(i,j)$ sono i pesi degli archi. Per i grafi di tolleranza non pesati, $w(i,j) = 1$ per ogni arco. La variazione totale misura la «rugosità» della funzione sul grafo: è piccola se f è quasi-costante sulle componenti connesse e grande se oscilla molto tra elementi adiacenti. Per i predicati vaghi, $TV_G(f) = \sum_{\{i,j\} \in E} |f(x_i) - f(x_j)|$ e la ε^* -coerenza equivale alla condizione $TV_G(f) \leq \varepsilon^* \cdot |E|$ nel caso di archi tutti di peso 1 (bound uniforme, non l'esatta condizione di ε^* -coerenza che richiede $|f(x_i) - f(x_j)| \leq \varepsilon^*$ per ogni singolo arco).

Il problema di minimizzazione della variazione totale con vincolo di contrasto è il problema di Rudin-Osher-Fatemi (ROF, 1992) discreto sul grafo di tolleranza:

$$f^* = \operatorname{argmin}_{\{f: X \rightarrow [0,1], f \text{ monotona}, f(x_1)=0, f(x_n)=1\}} TV_G(f)$$

Questo è un problema di programmazione lineare (la funzione obiettivo TV_G è lineare a tratti sui parametri f_i , e i vincoli sono lineari). La sua soluzione è un assegnamento a gradino (step function): la funzione ottimale fa un salto di ampiezza ε^* in ogni step del cammino monotono minimo e è

costante altrove. Questo recupera esattamente la costruzione dell'assegnamento ottimale f^* del Capitolo 5 da una prospettiva di ottimizzazione convessa della variazione totale.

Il collegamento con la metrica di Wasserstein è dato dal duale di Kantorovich: la variazione totale del predicato f è la norma duale W_1 della misura $(f - 1/2)$ rispetto alla distanza del grafo. Precisamente, per predicati «centrati» ($\int f d\mu = 1/2$):

$$TV_G(f) = \sup\{\sum_i \varphi_i (f_i - g_i) : \varphi \text{ 1-Lipschitz sul grafo } G\} = W_1(\mu_f, \mu_g)$$

per una scelta appropriata di g (il predicato «neutro» con $g_i = 1/2$ per ogni i). Questa identità stabilisce che la variazione totale semantica del predicato è la distanza di Wasserstein-1 dal predicato neutro: i predicati con alta variazione totale sono lontani dal neutro (hanno confini netti), quelli con bassa variazione totale sono vicini al neutro (massimamente vaghi).

Proposizione XIII.1 (Dualità tra vaghezza e costo di trasporto). Per ogni predicato $f \in \mathcal{F}_{\{\varepsilon^*\}}$ con $\bar{f} = 1/2$, la distanza di Wasserstein-1 dal predicato neutro $f_0 \equiv 1/2$ è: $W_1(\mu_f, \mu_{\{f_0\}}) = (1/n) \cdot TV_G(f) \cdot \text{diam}(G)/2$. Il numero di risoluzione ε^* è inversamente proporzionale al costo di trasporto minimo da f_0 al predicato ottimale f^* : $\varepsilon^*(\sim) = (TV_G(f^*)) / (|E| \cdot \text{diam}(G)) = 1/L_{\text{mono_min}}$.

XIII.4 L'equazione di continuità semantica e la geometria AB di Otto

Otto (2001) ha sviluppato una struttura riemanniana formale sullo spazio $P_2(M)$ delle misure di probabilità su una varietà M , mostrando che l'equazione di diffusione (l'equazione di Fokker-Planck) può essere interpretata come il flusso del gradiente dell'entropia di Boltzmann rispetto alla metrica di Wasserstein-2. Questa struttura — la geometria di Otto — si trasferisce allo spazio dei predicati vaghi.

La struttura di spazio metrico di $(\mathcal{F}_{\{\varepsilon^*\}}, W_2)$ si dota formalmente di una struttura riemanniana (lo spazio tangente in ogni punto $f \in \mathcal{F}_{\{\varepsilon^*\}}$ è lo spazio delle funzioni di potenziale discreto $\varphi: X \rightarrow \mathbb{R}$, con metrica indotta dal trasporto ottimale). La «derivata covariante» di una curva f_t in questo spazio è legata alla «velocità di trasporto» $v_t = \partial_t f_t$ tramite l'equazione di continuità semantica discreta:

$$\partial_t f_i + \sum_j : (i, j) \in E \ (v_{ij} \cdot f_i - v_{ji} \cdot f_j) = 0, \quad f_i(0) = f_0(x_i)$$

dove $v_{ij} \geq 0$ è il flusso di massa dall'elemento x_i verso x_j lungo l'arco (i, j) . L'equazione di continuità descrive la conservazione della massa totale: la perdita di massa in x_i è bilanciata dai flussi uscenti verso i vicini e dai flussi entranti dai vicini. Il costo totale del trasporto lungo la curva è:

$$\text{Action}(f_t, v_t) = \int_0^1 \sum_{\{(i, j) \in E\}} (v_{ij})^2 / f_i dt$$

e la distanza $W_2(f_0, f_1)$ è la radice quadrata del minimo dell'azione sulla curva che soddisfa l'equazione di continuità con condizioni iniziale e finale. Questo è il principio di minima azione semantica: la «traiettoria» ottimale tra due predicati vaghi f_0 e f_1 è quella che minimizza il costo cinetico $\sum_{\{(i, j) \in E\}} (v_{ij})^2 / f_i$ integrato nel tempo.

Nella geometria di Otto, l'equazione del calore semantica $\partial_t f = \Delta_{\sim} f$ è il flusso del gradiente dell'entropia di Gibbs $H(f) = -\sum_i f_i \log f_i$ rispetto alla metrica W_2 : i predicati evolvono lungo la direzione che massimizza l'aumento dell'entropia per unità di costo di trasporto. I punti stazionari del flusso sono i predicati che massimizzano l'entropia sotto il vincolo della struttura di tolleranza: questi sono i predicati con $f_i = \text{costante}$ (massima entropia dato il poliedro dei vincoli), cioè i predicati «massimalmente vaghi».

PARTE XIV. EQUAZIONI DIFFERENZIALI DISCRETE SUI GRAFI DI TOLLERANZA: PROPAGAZIONE E DIFFUSIONE DELLA VAGHEZZA

XIV.1 Il laplaciano del grafo di tolleranza e la sua interpretazione semantica

Il laplaciano del grafo $G(X, \sim)$ è l'operatore lineare $L: \mathbb{R}^s \rightarrow \mathbb{R}^s$ (con $s = |X|$) definito da:

$$(Lf)(x_i) = \sum_{j: (x_i, x_j) \in E} (f(x_i) - f(x_j)) = \deg(x_i) f(x_i) - \sum_{j: (x_i, x_j) \in E} f(x_j)$$

oppure in forma matriciale $L = D - A$, dove $D = \text{diag}(\deg(x_1), \dots, \deg(x_n))$ è la matrice dei gradi e A è la matrice di adiacenza del grafo. Il laplaciano normalizzato è $\mathcal{L} = D^{-1/2} L D^{-1/2} = I - D^{-1/2} A D^{-1/2}$.

L'interpretazione semantica del laplaciano è la seguente: per un predicato f , $(Lf)(x_i)$ misura quanto $f(x_i)$ si discosta dalla media dei valori di f negli elementi adiacenti a x_i nel grafo di tolleranza. Se f è coerente ($f(x_i) = f(x_j)$ per tutti i vicini x_j), allora $(Lf)(x_i) = 0$: il laplaciano è zero sui predicati coerenti. Se f ha una grande incoerenza locale in x_i ($f(x_i)$ è molto diverso dai vicini), allora $(Lf)(x_i)$ è grande in valore assoluto. Dunque il laplaciano misura la «curvatura semantica locale» del predicato: quanto il valore del predicato in x_i si discosta dal consenso locale.

Il quoziente di Rayleigh del laplaciano è:

$$R_L(f) = (f^T L f) / (f^T f) = \sum_{(i,j) \in E} (f(x_i) - f(x_j))^2 / \sum_i f(x_i)^2$$

e il minimo di $R_L(f)$ su tutti i vettori ortogonali al vettore costante è il primo autovalore non-nullo λ_1 (il gap spettrale). Il gap spettrale è il principale indicatore della «velocità di propagazione dell'informazione» nel grafo: grafi con λ_1 grande trasmettono informazione rapidamente (il

predicato si «uniforma» velocemente sotto diffusione), grafi con λ_1 piccolo trasmettono lentamente (la vaghezza persiste a lungo).

Per la catena lineare di n elementi con tolleranza uniforme di larghezza k , il gap spettrale è:

$$\lambda_1 = 2k(1 - \cos(\pi/(n-1))) \approx 2k\pi^2/(n-1)^2 \quad \text{per } n \gg 1$$

e la relazione con il numero di risoluzione $\varepsilon^* = k/(n-1)$ è:

$$\lambda_1 \approx 2\pi^2 \cdot \varepsilon^*(\sim) \cdot k/(n-1) = 2\pi^2 \cdot \varepsilon^*(\sim)^2 / (1/n)$$

Dunque il gap spettrale scala come il quadrato del numero di risoluzione per predicati unidimensionali: predicati più vaghi (ε^* piccolo) hanno gap spettrale più piccolo e diffondono l'informazione più lentamente.

XIV.2 L'equazione del calore semantica e la diffusione della vaghezza

L'equazione del calore sul grafo di tolleranza è l'equazione differenziale ordinaria:

$$df/dt = -L f(t), \quad f(0) = f_0$$

dove $f(t) = (f(x_1, t), \dots, f(x_n, t))$ è il vettore dei valori della funzione di membership al tempo t e L è il laplaciano del grafo. La soluzione è data dall'esponenziale di matrice:

$$f(t) = e^{-tL} f_0 = \sum_k e^{-\lambda_k t} \langle f_0, \psi_k \rangle \psi_k$$

dove $\lambda_0 = 0 \leq \lambda_1 \leq \dots \leq \lambda_{n-1}$ sono gli autovalori di L e $\psi_0, \psi_1, \dots, \psi_{n-1}$ le corrispondenti autofunzioni ortonormali. La componente $\lambda_0 = 0$ corrisponde all'autofunzione costante $\psi_0 = 1/\sqrt{n}$, e il contributo di ψ_0 a $f(t)$ è la proiezione di f_0 sulla costante: $\langle f_0, \psi_0 \rangle \psi_0 = \bar{f}_0 \cdot 1$ (la media). Per $t \rightarrow \infty$, tutti i termini con $\lambda_k > 0$ decadono esponenzialmente e la soluzione converge alla funzione costante $f_\infty = \bar{f}_0 \cdot 1$: il predicato diventa uniformemente uguale alla sua media iniziale.

Dal punto di vista della vaghezza, questa evoluzione ha un'interpretazione precisa: partendo da un predicato iniziale f_0 (non necessariamente coerente), l'equazione del calore lo evolve verso la coerenza perfetta (un predicato costante su ogni componente connessa). Il tasso di convergenza è determinato dal gap spettrale λ_1 : dopo un tempo $t \gg 1/\lambda_1$, il predicato è entro un errore $e^{-\lambda_1 t}$ dalla costante. Per predicati iniziali con contrasto ($f_0(x_1) = 0, f_0(x_n) = 1$), la diffusione «cancella» il contrasto in tempo $\tau_{\text{diff}} \approx \log(1/\varepsilon)/\lambda_1$.

Teorema XIV.1 (Tempo di diffusione e numero di risoluzione). Sia f_0 il predicato iniziale con $f_0(x_1)=0, f_0(x_n)=1$ e $f_0(x_i) = (i-1)/(n-1)$ (funzione lineare di Łukasiewicz). Dopo un tempo t , la deviazione del predicato diffuso $f(t) = e^{-tL} f_0$ dal predicato costante $f_\infty = 1/2$ è: $\|f(t) - f_\infty\|_2 \leq e^{-\lambda_1 t} \cdot \|f_0 - f_\infty\|_2 = e^{-\lambda_1 t} \cdot 1/\sqrt{12}$. Il tempo τ_{ε} tale che $\|f(\tau_{\varepsilon}) - f_\infty\|_2 \leq \varepsilon$ è: $\tau_{\varepsilon} = \log(1/(\varepsilon\sqrt{12})) / \lambda_1 \approx \log(1/\varepsilon) / (2\pi^2 \cdot \varepsilon^*(\sim)^2 \cdot n)$. Predicati più vaghi (ε^* piccolo) impiegano un tempo $\tau_{\varepsilon} \propto 1/\varepsilon^*(\sim)^2$ per perdere il contrasto — la vaghezza è 'più resistente' alla diffusione.

Dimostrazione. La funzione lineare $f_0(x_i) = (i-1)/(n-1)$ ha decomposizione spettrale $f_0 = \sum_k \langle f_0, \psi_k \rangle \psi_k$. Il termine dominante a grandi t è quello con il più piccolo autovalore λ_1 : $\|f(t) - f_0\|_2 \approx e^{-\lambda_1 t} \cdot |\langle f_0 - 1/2, \psi_1 \rangle| \leq e^{-\lambda_1 t} \cdot \|f_0 - 1/2\|_2 = e^{-\lambda_1 t} / \sqrt{12}$. Per la catena con tolleranza di larghezza k , $\lambda_1 \approx 2k\pi^2/(n-1)^2 = 2\pi^2 \epsilon^*(\sim)^2 \cdot (n-1)^{-2} \cdot (n-1)^2 \cdot 1 = 2\pi^2 \epsilon^*(\sim)^2$ quando normalizzato sulla dimensione del dominio. ■

XIV.3 L'equazione delle onde semantica e la propagazione dei confini del predicato

Accanto all'equazione del calore (di tipo parabolico), si considera l'equazione delle onde semantica (di tipo iperbolico) sul grafo di tolleranza:

$$d^2 f / dt^2 = -L f(t), \quad f(0) = f_0, \quad df/dt|_{t=0} = v_0$$

La soluzione è la superposizione di onde stazionarie:

$$f(t) = \sum_k [\cos(\sqrt{\lambda_k} t) \langle f_0, \psi_k \rangle + (\sin(\sqrt{\lambda_k} t) / \sqrt{\lambda_k}) \langle v_0, \psi_k \rangle] \psi_k$$

L'equazione delle onde descrive la propagazione di perturbazioni della funzione di membership: se si perturba localmente il predicato (un parlante cambia idea sull'elemento x_i), la perturbazione si propaga nel grafo con velocità determinata dagli autovalori. La velocità di propagazione delle onde nel grafo è $\text{conda} = \sqrt{\lambda_{n-1}}$ (*velocità massima, determinata dall'autovalore massimo*). Per la catena lineare con tolleranza uniforme di larghezza k , $\text{conda} = \sqrt{4k} = 2\sqrt{k}$.

L'interpretazione semantica dell'equazione delle onde è quella della «propagazione del confine del predicato»: se un gruppo di parlanti decide improvvisamente di spostare il confine semantico del predicato (cambio di norma linguistica), questa decisione si propaga nel resto della comunità linguistica con la velocità c_{onda} del grafo di tolleranza. Comunità con strutture di tolleranza più connesse (k grande) propagano i cambiamenti semantici più velocemente.

La dispersione dell'equazione delle onde sul grafo è non-nulla: le diverse componenti spettrali si propagano a velocità diverse ($\sqrt{\lambda_k}$ dipende da k), e una perturbazione inizialmente localizzata si «disperde» nel tempo. Il coefficiente di dispersione è la derivata seconda della relazione di dispersione $\omega(k) = \sqrt{\lambda_k}$ rispetto all'indice di modo k :

$$D_{\text{disp}} = \partial^2 \omega / \partial k^2 |_{k=1} = (1/2) (\lambda_2 - 2\lambda_1 + \lambda_0) / \lambda_1^{1/2}$$

Per la catena con tolleranza uniforme, $D_{\text{disp}} \approx -\pi^2/(n-1)^2 < 0$: il coefficiente di dispersione è negativo, indicando che le componenti ad alta frequenza si propagano più lentamente di quelle a bassa frequenza (dispersione «normale», analoga a quella di un medium dispersivo in ottica).

XIV.4 Il nucleo del calore semantico e i semigroup di diffusione

Il nucleo del calore (heat kernel) del grafo di tolleranza è la funzione:

$$K_{\text{t}}(x_i, x_j) = (e^{-tL})_{ij} = \sum_k e^{-\lambda_k t} \psi_k(x_i) \psi_k(x_j)$$

Il nucleo del calore ha le seguenti proprietà fondamentali: $Kt(x_i, x_j) \geq 0$ per ogni $t \geq 0$ (positività, dalla teoria dei semigroup di Markov); $\sum_j Kt(x_i, x_j) = 1$ per ogni i, t (conservazione della massa); $Kt(x_i, x_j) = Kt(x_j, x_i)$ (simmetria, poiché il grafo è non orientato e il laplaciano è simmetrico); $Ks(Kt(x, \cdot))(y) = K_{s+t}(x, y)$ (proprietà di semigroup). Per il grafo di tolleranza, il nucleo del calore ha un'espansione small-time:

$$K_t(x_i, x_j) = (1/n) [1 + e^{-\lambda_1 t} \psi_1(x_i) \psi_1(x_j) + O(e^{-\lambda_2 t})] \text{ per } t \rightarrow \infty$$

e un'espansione large-time (per $t \rightarrow 0$):

$$K_t(x_i, x_j) \approx (1/(4\pi t)^{d/2}) \exp(-d_G(x_i, x_j)^2/(4t)) \text{ per } t \rightarrow 0^+$$

dove d è la dimensione intrinseca del grafo (cioè la dimensione della varietà riemanniana da cui il grafo è campionato, secondo il Teorema di Latschev del Capitolo 3). L'espansione small-time è l'analoga discreta dell'espansione del nucleo del calore di Minakshisundaram-Pleijel per varietà riemanniane compatte, che lega il nucleo del calore al volume, alla curvatura scalare e agli invarianti spettrali della varietà.

La *distanza di diffusione* tra due elementi $x_i, x_j \in X$ è definita come:

$$d_{\text{diff}}(x_i, x_j; t) = \|K_t(x_i, \cdot) - K_t(x_j, \cdot)\|_2 = \sqrt{\sum_k (e^{-2\lambda_k t} (\psi_k(x_i) - \psi_k(x_j))^2)}$$

introdotta da Coifman e Lafon (2006) come strumento per la riduzione dimensionale e l'analisi dei dati. La distanza di diffusione generalizza la distanza del grafo d_G con maggiore robustezza al rumore: due elementi sono vicini rispetto a d_{diff} se sono connessi da molti cammini corti nel grafo (non solo da un unico cammino corto). Per il grafo di tolleranza, la distanza di diffusione misura la «vicinanza semantica tenendo conto di tutti i percorsi di indistinguibilità» tra due istanze del predicato, non soltanto del percorso più corto.

Proposizione XIV.1 (Distanza di diffusione e sorites barcode). Per il grafo di tolleranza $G(X, \sim)$ con filtrazione di Sorite parametrizzata da $r \in [0, \text{diam}(G)]$, la distanza di diffusione $d_{\text{diff}}(x_i, x_j; t)$ al tempo $t = 1/\lambda_1$ coincide, a meno di costanti universali, con la distanza di Gromov-Hausdorff della filtrazioni al livello $r = d_G(x_i, x_j)$. In particolare, gli elementi che compaiono nello stesso 'bar' del sorites barcode sono vicini rispetto a d_{diff} .

XIV.5 Equazioni di reazione-diffusione semantiche e formazione di pattern

Le equazioni di reazione-diffusione sul grafo di tolleranza combinano la diffusione laplaciana con un termine di «reazione» non lineare che modella l'effetto del contesto semantico:

$$df/dt = -L f + R(f), \quad f(0) = f_0$$

dove $R: \mathbb{R}^s \rightarrow \mathbb{R}^s$ è il termine di reazione semantica. Scelte naturali di R includono:

(a) Bistabilità (modello di Nagumo-Fisher): $R_i(f) = f_i(1-f_i)(f_i-\theta)$ dove $\theta \in (0,1)$ è la soglia di bistabilità. Questa equazione ha 2 stati stazionari stabili: $f \equiv 0$ (predicato sempre falso) e $f \equiv 1$

(predicato sempre vero), e uno stato instabile $f \equiv \theta$. La soluzione tende a «decidere» in ogni elemento x_i se va verso 0 o verso 1, producendo un pattern spaziale a gradino. La velocità con cui il fronte di separazione si propaga nel grafo è $c_{\text{fronte}} = 2\sqrt{(\lambda_1 \cdot \theta(1-\theta))}$.

(b) Inibizione laterale (modello di Turing): $R_i(f) = a_i f_i - b_i \sum_{j: x_i \sim x_j} f_j$, dove $a_i > 0$ è il coefficiente di auto-attivazione e $b_i > 0$ è il coefficiente di inibizione laterale. Se l'inibizione laterale è sufficientemente forte rispetto alla diffusione ($b_i/a_i > \lambda_1/\deg(x_i)$), il sistema sviluppa pattern di Turing: stati stazionari non omogenei in cui elementi alterni del grafo di tolleranza assumono valori alti e bassi della funzione di membership. Per i predicati vaghi, un pattern di Turing semantico corrisponderebbe a una struttura in cui elementi ordinalmente vicini (e quindi quasi indistinguibili) hanno valori di verità molto diversi — un fenomeno di «polarizzazione semantica» osservato in certi predicati politici.

(c) Potenziale di Ginzburg-Landau: $R_i(f) = -\partial V/\partial f_i$ con $V(f) = \sum_i (f_i^2/2 - f_i^4/4)$ (il doppio pozzo). Il potenziale di Ginzburg-Landau è il modello standard per le transizioni di fase continue: V ha 2 minimi in $f = \pm 1$ (corrispondenti ai valori netti Vero/Falso) e un massimo instabile in $f = 0$ (la zona di confine). L'equazione di Allen-Cahn sul grafo di tolleranza:

$$df/dt = -\varepsilon^{-2} \Delta f - \partial V/\partial f, \quad f(0) = f_0$$

nel limite $\varepsilon \rightarrow 0$ ha soluzioni che sviluppano interfacce nette di spessore $O(\varepsilon)$ che separano le regioni $f \approx +1$ (vero) e $f \approx -1$ (falso). La dinamica delle interfacce è governata dal moto per curvatura media sul grafo: la posizione dell'interfaccia si muove con velocità proporzionale alla curvatura locale della divisione vero/falso nella struttura del grafo di tolleranza. Il punto di equilibrio dell'interfaccia corrisponde alla «zona di confine ottimale» del predicato — la posizione in cui la curvatura è zero, che coincide con il punto critico del sorites barcode al livello di persistenza ε^* .

PARTE XV. LA TEORIA DELLE C*-ALGEBRE DEI PREDICATI VAGHI

XV.1 Motivazione e costruzione dell'algebra di tolleranza

Le C*-algebre sono la struttura algebrica fondamentale della meccanica quantistica non-commutativa e della geometria non-commutativa di Connes. La loro applicazione agli spazi di tolleranza è motivata dall'osservazione che le relazioni di tolleranza (non transitive) generano una struttura algebrica naturalmente non-commutativa: la composizione di indistinguibilità non è simmetrica rispetto all'ordine in cui si applicano le relazioni. Questo capitolo costruisce la C*-algebra associata a uno spazio di tolleranza, studia il suo spettro, e mostra come gli invarianti della C*-algebra catturino aspetti della struttura di vaghezza inaccessibili agli invarianti puramente combinatoriali o topologici.

Si considera lo spazio di Hilbert $\mathcal{H} = L^2(X, \mu) = \mathbb{C}^s$ (con $s = |X|$ e μ la misura uniforme su X). Ogni funzione $f: X \rightarrow \mathbb{C}$ agisce su $\mathcal{H}$ come operatore di moltiplicazione M_f : $(M_f \psi)(x) = f(x) \psi(x)$. Questi operatori formano una C*-algebra commutativa $\mathbb{C}^s$.

La C*-algebra di tolleranza è la C*-algebra $C^*(G(X, \sim))$ generata dagli operatori di moltiplicazione M_f (per $f: X \rightarrow \mathbb{C}$ arbitraria) e dagli operatori di «traslazione» $T_{\{ij\}}$ associati agli archi del grafo di tolleranza:

$$(T_{\{ij\}} \psi)(x_k) = \psi(x_k) + (\psi(x_j) - \psi(x_i)) \cdot \mathbb{1}_{\{k=i\}} - (\psi(x_j) - \psi(x_i)) \cdot \mathbb{1}_{\{k=j\}}$$

L'operatore $T_{\{ij\}}$ trasla massa (amplitude) dall'elemento x_i all'elemento x_j lungo l'arco (i, j) . Nel contesto semantico, $T_{\{ij\}}$ modella la «propagazione dell'indistinguibilità»: l'azione di $T_{\{ij\}}$ su un assegnamento ψ produce un nuovo assegnamento in cui l'informazione di x_i e x_j è uniformata. La C*-algebra generata da M_f e $T_{\{ij\}}$ al variare di f e $(i, j) \in E$ è la C*-algebra dell'algebra di tolleranza $C^*(X, \sim)$.

Definizione XV.1 (C*-algebra di tolleranza). Per uno spazio di tolleranza finito $(X, \sim)$, la C*-algebra di tolleranza è: $C^*(X, \sim) = C^*(M_f, T_{\{ij\}} : f \in C(X), (i, j) \in E(G(X, \sim))) \subset B(\mathcal{H})$, la sottoC*-algebra di $B(\mathcal{H}) = M_s(\mathbb{C})$ generata dagli operatori di moltiplicazione e traslazione. Per spazi di tolleranza finiti, $C^*(X, \sim)$ è una C*-algebra di dimensione finita, dunque isomorfa a una somma diretta di algebre di matrici: $C^*(X, \sim) \cong \bigoplus_k M_{n_k}(\mathbb{C})$.

XV.2 Il teorema di Gelfand-Naimark e lo spettro dell'algebra di tolleranza

Il teorema di Gelfand-Naimark (1943) stabilisce che ogni C^* -algebra commutativa è isomorfa (come C^* -algebra) all'algebra delle funzioni continue su un spazio compatto di Hausdorff — lo spettro dell'algebra. Per la C^* -algebra commutativa C^s delle funzioni su X , lo spettro è X stesso: i caratteri (omomorfismi moltiplicativi $C^s \rightarrow \mathbb{C}$) sono le valutazioni $\delta_x: f \mapsto f(x)$, uno per ogni $x \in X$. Per la C^* -algebra non-commutativa di tolleranza $C^*(X, \sim)$, lo spettro è più complicato.

Per C^* -algebre di dimensione finita, lo spettro (lo spazio dei caratteri irriducibili o rappresentazioni irriducibili) si calcola dalla decomposizione in somma diretta di algebre di matrici. Se $C^*(X, \sim) \cong \bigoplus_k M_{\{n_k\}}(\mathbb{C})$, le rappresentazioni irriducibili sono le proiezioni sui blocchi $M_{\{n_k\}}(\mathbb{C})$, di dimensione n_k . Le dimensioni n_k soddisfano la relazione:

$$\dim C^*(X, \sim) = \sum_k n_k^2 \leq s^2$$

con uguaglianza se e solo se $C^*(X, \sim) = M_s(\mathbb{C})$ (il caso in cui tutti gli elementi di X sono collegati nel grafo di tolleranza, cioè il grafo è completo). La struttura della decomposizione in blocchi è strettamente legata alla struttura del grafo di tolleranza:

Teorema XV.1 (Struttura della C^* -algebra di tolleranza). Sia $(X, \sim)$ uno spazio di tolleranza con grafo $G(X, \sim)$ avente componenti connesse $C_1, \dots, C_k$. Allora: (i) $C^*(X, \sim) \cong \bigoplus_i C^*(C_i, \sim|_{C_i})$ (decomposizione sulle componenti connesse); (ii) Per ogni componente C_i connessa, $C^*(C_i, \sim|_{C_i}) \cong M_{|C_i|}(\mathbb{C})$ se il grafo di tolleranza ristretto a C_i è un grafo completo $K_{|C_i|}$; (iii) In generale, $C^*(C_i, \sim|_{C_i})$ ha dimensione $\dim C^*(C_i) \in [|C_i|, |C_i|^2]$ determinata dalla struttura del grafo ristretto. (iv) Il numero di classi di equivalenza di rappresentazioni irriducibili di $C^*(X, \sim)$ è uguale al numero di componenti connesse k del grafo $G(X, \sim)$.

Dimostrazione (sketch). (i) Segue dal fatto che gli operatori $T_{\{ij\}}$ con (i, j) attraverso due componenti diverse sono zero (non ci sono archi tra componenti diverse). Dunque la C^* -algebra si decompone come somma diretta sulle componenti. (ii) Per il grafo completo K_m , ogni coppia (i, j) ha un arco. Gli operatori $T_{\{ij\}}$ generano tutte le matrici elementari E_{ij} (basi delle matrici $m \times m$): $T_{\{ij\}} - T_{\{ji\}} = E_{ii} - E_{jj}$ (differenza di proiettori), e $(T_{\{ij\}} - T_{\{ji\}}) \cdot (T_{\{jk\}} - T_{\{kj\}}) = E_{ik}$ (prodotto di matrici elementari). Dunque $C^*(K_m) = M_m(\mathbb{C})$. (iii) Segue dalla struttura generale delle C^* -algebre di grafi (Cuntz-Krieger). (iv) Le rappresentazioni irriducibili di $\bigoplus_i M_{\{n_i\}}(\mathbb{C})$ sono esattamente k , una per componente. ■

XV.3 Stati KMS e termodinamica della vaghezza

Nella meccanica quantistica statistica, gli stati di equilibrio termodinamico a temperatura inversa β sono gli stati KMS (Kubo-Martin-Schwinger): stati ω sulla C^* -algebra dell'osservabile del sistema che soddisfano la condizione di stazionarietà della funzione di correlazione nel senso analitico. Per

la C*-algebra di tolleranza $C^*(X, \sim)$, i corrispondenti stati KMS modellano l'«equilibrio semantico» di un parlante che assegna valori di verità al predicato.

Si introduce il gruppo di automorfismi di $C^*(X, \sim)$ corrispondente all'evoluzione temporale semantica: il gruppo $\{\alpha^t\}_{t \in \mathbb{R}}$ dove $\alpha^t(A) = e^{\{itH\}} A e^{\{-itH\}}$ e $H = L$ è il laplaciano del grafo di tolleranza (l'hamiltoniana semantica). Lo stato β -KMS ω_β sulla C*-algebra $C^*(X, \sim)$ soddisfa:

$$\omega_\beta(A \cdot \alpha^t(B)) = \omega_\beta(\alpha^t(B) \cdot A), \quad \text{per } A, B \in C^*(X, \sim) \text{ e } t \text{ esteso nel semipiano complesso } t + i\beta$$

La soluzione esplicita dello stato β -KMS per l'hamiltoniana L è lo stato di Gibbs:

$$\omega_\beta(A) = \text{Tr}(A \cdot e^{\{-\beta L\}}) / \text{Tr}(e^{\{-\beta L\}}) = \sum_k \langle \psi_k, A \psi_k \rangle e^{\{-\beta \lambda_k\}} / Z(\beta)$$

dove $Z(\beta) = \text{Tr}(e^{\{-\beta L\}}) = \sum_k e^{\{-\beta \lambda_k\}}$ è la funzione di partizione. Per $\beta \rightarrow 0$ (temperatura alta), lo stato KMS tende alla traccia normalizzata (equidistribuzione sulle autofunzioni); per $\beta \rightarrow \infty$ (temperatura zero), lo stato KMS si concentra sull'autospazio del minimo autovalore $\lambda_0 = 0$ (lo stato fondamentale, corrispondente alla funzione costante).

La funzione di partizione $Z(\beta)$ e la funzione energia libera $F(\beta) = -(1/\beta) \log Z(\beta)$ hanno espressioni esplicite per il grafo di tolleranza in funzione degli autovalori:

$$Z(\beta) = \sum_{k=0}^{\infty} e^{\{-\beta \lambda_k\}}, \quad F(\beta) = -(1/\beta) \log(\sum_k e^{\{-\beta \lambda_k\}})$$

L'entropia termodinamica (von Neumann) dello stato KMS è:

$$S(\beta) = -d(\beta F)/d\beta = \beta^2 dF/d\beta = \log Z(\beta) - \beta \langle L \rangle_\beta$$

dove $\langle L \rangle_\beta = \sum_k \lambda_k e^{\{-\beta \lambda_k\}}/Z(\beta)$ è l'energia media. Per $\beta \rightarrow 0$, $S(\beta) \rightarrow \log s$ (entropia massima, tutte le autofunzioni equiprobabili); per $\beta \rightarrow \infty$, $S(\beta) \rightarrow \log(\dim \ker L)$ (entropia del ground state, pari al logaritmo del numero di componenti connesse k).

Teorema XV.2 (Transizione di fase semantica nell'algebra di tolleranza). Sia $C^*(X, \sim)$ la C*-algebra di tolleranza con hamiltoniana L avente gap spettrale $\lambda_1 > 0$. Esiste una temperatura critica $\beta_c = 1/\lambda_1$ tale che: (i) Per $\beta < \beta_c$ (temperatura alta): lo stato KMS ω_β è unico e mischia tutte le componenti spettrali; (ii) Per $\beta > \beta_c$ (temperatura bassa): lo stato KMS ω_β è dominato dalla componente $\lambda_0 = 0$ (funzione costante) con correzioni di ordine $e^{\{-\beta \lambda_1\}}$; (iii) La capacità calorifica $C(\beta) = \beta^2 d^2F/d\beta^2$ ha un picco a $\beta \approx \beta_c$ con $C(\beta_c) \approx (\lambda_1 - \lambda_0)^2 \cdot n/4$ (n = numero di autovalori nel gap $\lambda_0 < \lambda < \lambda_1$, che per grafi di tolleranza connessi è 0, dunque il picco ha larghezza $\propto (\lambda_2 - \lambda_1)^{-1}$).

XV.4 La K-teoria della C*-algebra di tolleranza e gli invarianti semantici

La K-teoria delle C*-algebre (Atiyah, 1964; Hirzebruch, 1956; Bott, 1959) classifica le C*-algebre tramite due gruppi abeliani: K_0 (che classifica le proiezioni, ossia le rappresentazioni irriducibili, a meno di equivalenza) e K_1 (che classifica le unitarie). Per C*-algebre di dimensione finita, la K-teoria è particolarmente accessibile: per $M_n(\mathbb{C})$, $K_0(M_n(\mathbb{C})) \cong \mathbb{Z}$ (generato dalla proiezione di rango

1) e $K_1(M_n(\mathbb{C})) = 0$ (tutte le unitarie sono connesse all'identità). Per la C^* -algebra di tolleranza $C^*(X, \sim) \cong \bigoplus_k M\{n_k\}(\mathbb{C})$:

$$K_0(C^*(X, \sim)) \cong \mathbb{Z}^k, \quad K_1(C^*(X, \sim)) = 0$$

dove k è il numero di blocchi nella decomposizione — cioè il numero di componenti connesse di $G(X, \sim)$, come stabilito dal Teorema XV.1. Il risultato $K_0 \cong \mathbb{Z}^k$ è l'analogo K-teorico della Proposizione 1 del Capitolo 2: il numero di assegnamenti coerenti indipendenti è 2^k , e il gruppo K_0 ha rango k .

La K-teoria diventa più interessante per C^* -algebre associate a spazi di tolleranza infiniti (ad esempio, lo spazio delle stature come sottoinsieme denso di $\mathbb{R}$). In questo caso, la C^* -algebra di tolleranza è la C^* -algebra del groupoide dell'olonomia del grafo di tolleranza — un oggetto non-commutativo di dimensione infinita la cui K-teoria cattura invarianti topologici del grafo a scala infinita.

Per spazi di tolleranza metrici (X, d_X) con la tolleranza $\sim_\varepsilon$ per un parametro ε fissato, la C^* -algebra di tolleranza è la C^* -algebra del groupoide di tolleranza $G\{\varepsilon\}(X) = \{(x, y) \in X \times X : d_X(x, y) \leq \varepsilon\}$ con la struttura di groupoide dove la composizione $(x, y) \cdot (y, z) = (x, z)$ è definita quando $d_X(x, y) \leq \varepsilon$ e $d_X(y, z) \leq \varepsilon$ (ma non necessariamente $d_X(x, z) \leq \varepsilon$ — la non-transitività appare anche nella struttura del groupoide!). La C^* -algebra del groupoide $C^*(G_\varepsilon(X))$ è la versione non-commutativa dell'algebra delle funzioni continue sull'insieme delle coppie ε -vicine, e cattura la geometria non-commutativa della struttura di indistinguibilità.

Teorema XV.3 (Connessione tra K-teoria e omologia persistente). Per la famiglia parametrizzata di C^* -algebre di tolleranza $\{C^*(G_{\varepsilon}(X))\}_{\varepsilon \geq 0}$ (la filtrazione di C^* -algebre), il mappa di inclusione $C^*(G_{\varepsilon}(X)) \rightarrow C^*(G_{\varepsilon'}(X))$ per $\varepsilon \leq \varepsilon'$ induce mappe $K_0(C^*(G_{\varepsilon}(X))) \rightarrow K_0(C^*(G_{\varepsilon'}(X)))$. I moduli persistenti associati a queste mappe in K_0 sono isomorfi ai moduli persistenti di H_0 del complesso di Sorite: il sorites barcode in dimensione 0 coincide con il 'barcode K_0 ' della filtrazione di C^* -algebre. Per le dimensioni superiori (K_1 e K-teoria di ordine superiore), il barcode K_1 codifica informazioni topologiche che corrispondono agli H_1 e H_2 del complesso di Sorite.

XV.5 La geometria non-commutativa e il triplo spettrale della vaghezza

La geometria non-commutativa di Alain Connes (1994) descrive gli spazi geometrici non tramite punti e funzioni ma tramite tripli spettrali $(A, \mathcal{H}, D)$: una C^* -algebra A , uno spazio di Hilbert $\mathcal{H}$ su cui A agisce, e un operatore di Dirac D che codifica la struttura metrica. Per gli spazi di tolleranza, il triplo spettrale naturale è:

$$A = C^*(X, \sim), \quad \mathcal{H} = L^2(X) \oplus L^2(E), \quad D = \begin{pmatrix} 0 & d^* \\ d & 0 \end{pmatrix}$$

dove $d: L^2(X) \rightarrow L^2(E)$ è l'operatore di bordo del grafo (il cobord) e $d^*: L^2(E) \rightarrow L^2(X)$ è il suo aggiunto. L'operatore D è l'operatore di Dirac del grafo di tolleranza: $D^2 = d^*d \oplus dd^* = (L \oplus \tilde{L})$

dove L è il laplaciano sui vertici e $\tilde{L} = dd^*$ è il laplaciano sugli archi. La metrica di Connes indotta dal triplo spettrale sulla stato puro (un punto $x \in X$) è:

$$d_C(x, y) = \sup\{|f(x) - f(y)| : f \in A, \|[D, Mf]\|_{B(\mathcal{H})} \leq 1\}$$

dove $[D, Mf] = D Mf - Mf D$ è il commutatore di D con l'operatore di moltiplicazione Mf . Il calcolo del commutatore dà $[D, Mf] \psi = d(f\psi) - f d(\psi) = (df)\psi$, dove $df: E \rightarrow \mathbb{C}$ è la differenza prima della funzione f sugli archi: $(df)(i, j) = f(x_j) - f(x_i)$. Dunque $\|[D, Mf]\| = \|df\|_\infty = \max\{|f(x_j) - f(x_i)| : (i, j) \in E\}$ e la condizione $\|[D, Mf]\| \leq 1$ equivale a $|f(x_j) - f(x_i)| \leq 1$ per ogni arco (i, j) del grafo di tolleranza. La metrica di Connes è dunque:

$$d_C(x, y) = \sup\{|f(x) - f(y)| : |f(x_i) - f(x_j)| \leq 1 \text{ per ogni } (i, j) \in E\}$$

che è esattamente la distanza del grafo $d_G(x, y)$ (numero minimo di archi in un cammino da x a y)! Questo risultato — la coincidenza tra la metrica di Connes e la distanza del grafo — mostra che il triplo spettrale della vaghezza recupera esattamente la geometria discreta del grafo di tolleranza, ricollocandola nel framework della geometria non-commutativa.

Per il numero di risoluzione ε^* , la formulazione tramite la metrica di Connes è:

$$\begin{aligned} \varepsilon^*(\sim) &= 1 / \max\{|f(x_n) - f(x_1)| : f \in A, \|[D, Mf]\| \leq 1, f \text{ monotona}\} \\ &= 1 / d_G^{\{\text{mono}\}}(x_1, x_n) = 1 / L_{\text{mono_min}} \end{aligned}$$

dove $d_G^{\{\text{mono}\}}$ è la distanza monotona nel DAG $G^{\rightarrow}(X, \sim)$ — la lunghezza del cammino monotono minimo. Il triplo spettrale della vaghezza codifica dunque tutta la struttura del numero di risoluzione ε^* nella geometria non-commutativa dell'operatore di Dirac D .

XV.6 Il teorema dell'indice di Atiyah-Singer per i predicati vaghi

Il teorema dell'indice di Atiyah-Singer (1963, 1968) è uno dei risultati più profondi della matematica del XX secolo: esso afferma che l'indice analitico di un operatore ellittico su una varietà compatta (definito come $\dim \ker D - \dim \text{coker } D$) coincide con l'indice topologico (calcolato come integrale di certe classi caratteristiche della varietà). Per il triplo spettrale della vaghezza, il teorema dell'indice prende una forma discreta particolarmente esplicita.

L'operatore D del triplo spettrale della vaghezza è l'operatore di Dirac del grafo, e il suo indice (nel senso del grafo) è:

$$\text{ind}(D) = \dim \ker D_+ - \dim \ker D_-$$

dove $D_+ = d: L^2(X) \rightarrow L^2(E)$ è il cobord e $D_- = d^*: L^2(E) \rightarrow L^2(X)$ è il bordo. Il nucleo di $D_+ = d$ è $\ker d = \{f: X \rightarrow \mathbb{C} : f \text{ costante su ogni componente connessa}\}$ (le funzioni localmente costanti, cioè i cicli 0-dimensionali), con dimensione k (il numero di componenti). Il conucleo di D_+ è l'immagine di d^* meno l'intera $L^2(E)$: $\dim \text{coker } D_+ = \dim L^2(E) - \dim \text{im } d = |E| - (|X| - k) = \mu(G) = |E| - |X| + k$ (il numero ciclomático). Dunque:

$$\text{ind}(D_+) = \dim \ker D_+ - \dim \text{coker } D_+ = k - (|E| - |X| + k) = |X| - |E|$$

che è la caratteristica di Euler-Poincaré del grafo $\chi(G) = |X| - |E|$ (con tutti i triangoli e semplici superiori contati con segno nella formula di Euler, qui applicata al caso 1-dimensionale). L'identità $\text{ind}(D_+) = \chi(G)$ è la versione discreta del teorema dell'indice di Atiyah-Singer: l'indice analitico dell'operatore di Dirac del grafo coincide con l'invariante topologico $\chi(G)$. Il lato topologico del teorema si calcola tramite la formula di Gauss-Bonnet per grafi:

$$\chi(G) = \sum_v \kappa(v) = \sum_v (1 - \deg(v)/2) \quad (\text{formula di Gauss-Bonnet discreta, per grafi 2-cell embedded})$$

dove $\kappa(v) = 1 - \deg(v)/2$ è la curvatura discreta in un vertice v (analoga alla curvatura gaussiana per una superficie). La formula afferma che la caratteristica di Euler è la somma delle curvature discrete sui vertici — l'analogo del teorema di Gauss-Bonnet che esprime χ come integrale della curvatura gaussiana.

Corollario XV.1 (Teorema dell'indice per la vaghezza). Per uno spazio di tolleranza $(X, \sim)$ con grafo $G(X, \sim)$ connesso, l'indice dell'operatore di Dirac della vaghezza soddisfa: $\text{ind}(D_{\text{vaghezza}}) = \chi(G(X, \sim)) = |X| - |E(G)| = 1 - \mu(G(X, \sim))$, dove $\mu(G) = |E| - |X| + 1$ è il numero ciclomatico (il numero di ostruzioni omologiche indipendenti del Capitolo 6). Il numero di ostruzioni omologiche è dunque $\mu = 1 - \text{ind}(D_{\text{vaghezza}})$: le ostruzioni omologiche all'assegnamento coerente sono contate dall'indice meno 1 dell'operatore di Dirac della vaghezza.

Questo corollario chiude il cerchio tra la topologia combinatoria (il numero ciclomatico μ , il numero di ostruzioni omologiche ai predicati coerentemente non-costanti), la teoria degli operatori (l'indice dell'operatore di Dirac del grafo) e la geometria non-commutativa (il triplo spettrale della vaghezza). Il paradosso di Sorite, descritto nel Capitolo 1 come antinomia semantica, si traduce in questo framework nella formula $\text{ind}(D) = 1 - \mu$: se il numero di ostruzioni omologiche $\mu \geq 1$ (e per grafi connessi con cicli, $\mu \geq 0$, con $\mu = 0$ se e solo se il grafo è un albero), l'indice dell'operatore di Dirac è ≤ 1 . Per grafi connessi non-ciclici (alberi), $\mu = 0$ e $\text{ind} = 1$: l'operatore di Dirac ha un unico «modo zero», corrispondente all'unico assegnamento coerente a contrasto (la funzione che vale 0 alla radice dell'albero e cresce verso le foglie). Per grafi connessi con cicli, $\mu \geq 1$ e $\text{ind} \leq 0$: l'operatore di Dirac ha più «modi zero nel cokernel» che nel kernel — le ostruzioni omologiche superano le soluzioni, e il paradosso di Sorite è più profondo.

TAVOLA RIASSUNTIVA: I DIZIONARI DI CORRISPONDENZA DELLE PARTI XII–XV

Le quattro sezioni sviluppate in questo supplemento stabiliscono ulteriori ponti sistematici tra la teoria degli spazi di tolleranza e quattro domini della matematica contemporanea. Si riassumono le corrispondenze principali, con indicazione del grado di precisione e dei punti aperti.

Parte	Territorio matematico → Spazio di tolleranza	Risultato principale
XII	Metrica di Fisher → Misura di distinguibilità tra predicati	Bound di Cramér-Rao: la zona di confine richiede $O(1/\delta^2)$ classificazioni per stima con errore δ
XII	Entropia di Fisher → Vaghezza massima	$S_F(f^*)$ massima per $f^* \equiv 1/2$ (predicato massimalmente vago)
XIII	Metrica W_1 → Variazione totale semantica	$TV_G(f) = W_1(\mu_f, \mu_{f_0})$: la variazione totale è la distanza Wasserstein dal predicato neutro
XIII	Geodetiche di McCann → Trasporto tra predicati	Contrazione: $W_2(e^{t\Delta}f, e^{t\Delta}g) \leq e^{-\lambda_1 t} W_2(f, g)$ con tasso λ_1
XIV	Equazione del calore → Diffusione della vaghezza	$\tau_\varepsilon \approx \log(1/\varepsilon)/(2\pi^2 \varepsilon^{*2} \cdot n)$: predicati più vaghi resistono alla diffusione più a lungo
XIV	Equazione di Allen-Cahn → Formazione della zona di confine	Il fronte di separazione vero/falso si stabilizza sulla zona di confine ottimale del sorites barcode
XV	C^* -algebra di tolleranza → Struttura algebrica della vaghezza	$K_0(C^*(X, \sim)) \cong \mathbb{Z}^k$: il rango K_0 conta le componenti connesse (assegnamenti indipendenti)
XV	Metrica di Connes → Distanza del grafo	$d_C(x, y) = d_G(x, y)$: la geometria non-commutativa recupera la metrica del grafo di tolleranza
XV	Teorema dell'indice di Atiyah-Singer → Numero di ostruzioni	$\text{ind}(D_{\text{vaghezza}}) = 1 - \mu(G)$: le ostruzioni omologiche contano il difetto d'indice dell'operatore di Dirac

RIFERIMENTI AGGIUNTIVI PER LE SEZIONI XII–XV

- Amari S. e Nagaoka H., *Methods of Information Geometry*, American Mathematical Society & Oxford University Press (2000).
- Rao C. R., «Information and the accuracy attainable in the estimation of statistical parameters», *Bulletin of the Calcutta Mathematical Society* 37, 81–91 (1945).
- Cramér H., *Mathematical Methods of Statistics*, Princeton University Press (1946).
- Kullback S. e Leibler R. A., «On information and sufficiency», *Annals of Mathematical Statistics* 22(1), 79–86 (1951).
- Jaynes E. T., «Information theory and statistical mechanics», *Physical Review* 106, 620–630 (1957).
- Monge G., «Mémoire sur la théorie des déblais et remblais», *Histoire de l'Académie Royale des Sciences* (1781).
- Kantorovich L. V., «On the translocation of masses», *Doklady Akademii Nauk SSSR* 37(7–8), 227–229 (1942).
- Villani C., *Topics in Optimal Transportation*, Graduate Studies in Mathematics 58, American Mathematical Society (2003).
- Villani C., *Optimal Transport: Old and New*, Grundlehren der mathematischen Wissenschaften 338, Springer (2009).
- Otto F., «The geometry of dissipative evolution equations: the porous medium equation», *Communications in Partial Differential Equations* 26(1–2), 101–174 (2001).
- McCann R. J., «A convexity principle for interacting gases», *Advances in Mathematics* 128(1), 153–179 (1997).
- Brenier Y., «Polar factorization and monotone rearrangement of vector-valued functions», *Communications on Pure and Applied Mathematics* 44(4), 375–417 (1991).
- Benamou J.-D. e Brenier Y., «A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem», *Numerische Mathematik* 84(3), 375–393 (2000).
- Erbar M. e Maas J., «Ricci curvature of finite Markov chains via convexity of the entropy», *Archive for Rational Mechanics and Analysis* 206(3), 997–1038 (2012).
- Mielke A., «Geodesic convexity of the relative entropy in reversible Markov chains», *Calculus of Variations and Partial Differential Equations* 48(1), 1–31 (2013).
- Rudin L. I., Osher S. e Fatemi E., «Nonlinear total variation based noise removal algorithms», *Physica D: Nonlinear Phenomena* 60(1–4), 259–268 (1992).
- Coifman R. R. e Lafon S., «Diffusion maps», *Applied and Computational Harmonic Analysis* 21(1), 5–30 (2006).
- Nagumo J., Arimoto S. e Yoshizawa S., «An active pulse transmission line simulating nerve axon», *Proceedings of the IRE* 50(10), 2061–2070 (1962).
- Allen S. M. e Cahn J. W., «A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening», *Acta Metallurgica* 27(6), 1085–1095 (1979).
- Turing A. M., «The chemical basis of morphogenesis», *Philosophical Transactions of the Royal Society B* 237(641), 37–72 (1952).
- Connes A., *Noncommutative Geometry*, Academic Press (1994).
- Gelfand I. M. e Naimark M. A., «On the imbedding of normed rings into the ring of operators in Hilbert space», *Matematicheskii Sbornik* 12(54), 197–213 (1943).
- Atiyah M. F. e Singer I. M., «The index of elliptic operators on compact manifolds», *Bulletin of the American Mathematical Society* 69(3), 422–433 (1963).
- Atiyah M. F. e Singer I. M., «The index of elliptic operators: III», *Annals of Mathematics* 87(3), 546–604 (1968).
- Cuntz J. e Krieger W., «A class of C^* -algebras and topological Markov chains», *Inventiones Mathematicae* 56(3), 251–268 (1980).
- Blackadar B., *K-theory for Operator Algebras*, 2^a ed., Mathematical Sciences Research Institute Publications 5, Cambridge University Press (1998).
- Connes A., «On the spectral characterization of manifolds», *Journal of Noncommutative Geometry* 7(1), 1–82 (2013).
- Roe J., *Lectures on Coarse Geometry*, University Lecture Series 31, American Mathematical Society (2003).
- Higson N. e Roe J., *Analytic K-homology*, Oxford Mathematical Monographs, Oxford University Press (2000).
- Bakry D. e Émery M., «Diffusions hypercontractives», in *Séminaire de Probabilités XIX*, Lecture Notes in Mathematics 1123, 177–206, Springer (1985).
- Lott J. e Villani C., «Ricci curvature for metric-measure spaces via optimal transport», *Annals of Mathematics* 169(3), 903–991 (2009).

- Sturm K.-T., «On the geometry of metric measure spaces I e II», *Acta Mathematica* 196(1–2), 65–177 (2006).
- Gigli N., «On the differential structure of metric measure spaces and applications», *Memoirs of the American Mathematical Society* 236(1113) (2015).
- Gauss C. F., *Disquisitiones generales circa superficies curvas*, *Commentationes Societatis Regiae Scientiarum Gottingensis* (1828).
- Bonnet P.-O., «Mémoire sur la théorie générale des surfaces», *Journal de l'École Polytechnique* 32, 1–146 (1848).
- Minakshisundaram S. e Pleijel Å., «Some properties of the eigenfunctions of the Laplace-operator on Riemannian manifolds», *Canadian Journal of Mathematics* 1(3), 242–256 (1949).

PARTE XVI. GEOMETRIA ALGEBRICA TROPICALE E STRUTTURE DI TOLLERANZA: UN SEMINARIO SULLA VAGHEZZA SUL SEMIANELLO MAX-PLUS

XVI.1 Il semianello max-plus e la sua interpretazione semantica

La geometria tropicale, sviluppata sistematicamente da Mikhalkin (2003, 2005), Speyer e Sturmfels (2004), e Itenberg, Mikhalkin e Shustin (2007), studia le varietà algebriche su un campo non archimedeo mediante la sostituzione delle operazioni aritmetiche usuali (addizione e moltiplicazione su $\mathbb{R}$ o $\mathbb{C}$) con le operazioni del semianello max-plus: «addizione» $a \oplus b = \max(a, b)$ e «moltiplicazione» $a \otimes b = a + b$, con l'elemento neutro additivo $-\infty$ (identità per il massimo) e l'elemento neutro moltiplicativo 0 (identità per la somma). La struttura $(\mathbb{R} \cup \{-\infty\}, \oplus, \otimes)$ è un semianello (gli assiomi dei semianelli sono verificati, ma manca l'opposto additivo: $a \oplus (-a) \neq -\infty$ in generale) che prende il nome di semianello tropicale T.

La connessione con la teoria degli spazi di tolleranza è motivata da un'osservazione fondamentale: il numero di risoluzione $\varepsilon^*(\sim) = 1/L_{\text{mono_min}}$ (Teorema 3, Capitolo 5) si esprime come soluzione di un problema di ottimizzazione max-plus. Precisamente, la lunghezza del cammino monotono minimo $L_{\text{mono_min}}$ è la soluzione del problema di «moltiplicazione di matrici tropicale» sul DAG $G^{\rightarrow}(X, \sim)$. Questo non è una coincidenza formale, ma riflette il fatto che i problemi di cammino minimo su grafi — la struttura fondamentale del calcolo di ε^* — sono naturalmente descritti dall'algebra max-plus.

Si definisce la matrice di adiacenza tropicale del DAG $G^{\rightarrow}(X, \sim)$ come la matrice $A_{\text{trop}} \in T^{\wedge}\{n \times n\}$ con $(A_{\text{trop}})_{ij} = 0$ se esiste l'arco orientato $x_i \rightarrow x_j$ in $G^{\rightarrow}(X, \sim)$ (cioè se $x_i \sim x_j$ e $x_i < x_j$), e $(A_{\text{trop}})_{ij} = -\infty$ altrimenti. La moltiplicazione di matrici tropicale è definita dal prodotto scalare con $\oplus$ e $\otimes$:

$$(A \otimes_{\text{trop}} B)_{ij} = \max_k \{ A_{ik} + B_{kj} \}$$

La potenza tropicale $A_{\text{trop}}^{\{\otimes k\}}$ ha l'entrata $(A_{\text{trop}}^{\{\otimes k\}})_{ij}$ uguale alla lunghezza massima di un cammino orientato di esattamente k passi da x_i a x_j nel DAG (con $-\infty$ se non esiste tale cammino). Dualmente, per il problema di cammino minimo, si utilizza il semianello min-plus ($a \oplus b = \min(a,b)$, $a \otimes b = a+b$) — la versione «duale» — dove la matrice di Floyd-Warshall è la potenza min-plus iterata della matrice dei pesi:

$$D = A_{\text{minplus}}^{\{\otimes (n-1)\}} \quad (\text{chiusura di Kleene min-plus})$$

con $D_{\{1n\}} = L_{\text{mono_min}}$ e $\varepsilon^*(\sim) = 1/D_{\{1n\}}$. Il calcolo di ε^* è dunque la chiusura di Kleene nel semianello min-plus della matrice di adiacenza orientata: una delle operazioni fondamentali dell'algebra tropicale.

Definizione XVI.1 (Varietà tropicale di tolleranza). Per uno spazio di tolleranza $(X, \sim, \leq)$ con n elementi, la varietà tropicale di tolleranza $V_{\text{trop}}(X, \sim) \subset T^n$ è l'insieme delle funzioni di membership tropicali: le funzioni $f: X \rightarrow T = \mathbb{R} \cup \{-\infty\}$ che soddisfano la condizione di coerenza tropicale $|_{\text{trop}} f(x_i) \ominus_{\text{trop}} f(x_j)| \leq 0$ per ogni coppia tollerante $x_i \sim x_j$, dove la differenza tropicale è $a \ominus_{\text{trop}} b = a - b$ (la sottrazione reale, l'unica operazione invertibile nel semianello tropicale). Esplicitamente: $f \in V_{\text{trop}}(X, \sim)$ se e solo se $\max_{\{x_i \sim x_j\}} |f(x_i) - f(x_j)| \leq 0$, ovvero $f(x_i) = f(x_j)$ per ogni coppia tollerante.

La varietà tropicale di tolleranza è l'immagine «tropicale» della condizione di 0-coerenza: una funzione è nella varietà tropicale di tolleranza se e solo se è un assegnamento coerente nel senso classico (Definizione 8). Il Teorema 1 (Il paradosso di Sorite come teorema) si riformula in questo linguaggio: la varietà tropicale di tolleranza $V_{\text{trop}}(X, \sim)$ non contiene nessuna funzione con contrasto ($f(x_1) = -\infty$, $f(x_n) = 0$ nella parametrizzazione tropicale) quando $G(X, \sim)$ è connesso.

XVI.2 Polinomi tropicali e il polinomio caratteristico del grafo di tolleranza

Un polinomio tropicale è un'espressione della forma $p_{\text{trop}}(x) = \max_{\alpha} (c_{\alpha} + \sum_i \alpha_i x_i)$ dove $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n$ sono i monomi (vettori di esponenti interi) e $c_{\alpha} \in \mathbb{R}$ sono i coefficienti. Il «luogo degli zeri» di un polinomio tropicale è la sua ipersuperficie tropicale: l'insieme dei punti x dove il massimo è raggiunto da almeno 2 monomi simultaneamente (la condizione di «non-unicità del massimo» è l'analogo tropicale dello zero ordinario).

Il polinomio caratteristico tropicale del laplaciano L del grafo di tolleranza è il determinante tropicale:

$$P_{\text{trop}}(\lambda) = \det_{\text{trop}}(\lambda I - L) = \max_{\{\sigma \in S_n\}} \{ \text{sgn}_{\text{trop}}(\sigma) + \sum_i (\lambda I - L)_{i, \sigma(i)} \}$$

dove il determinante tropicale è definito come il massimo permanente sul gruppo simmetrico S_n (con segno tropicale: $+0$ per permutazioni pari e $-\infty$ per permutazioni dispari nel semianello con

segno). Per i grafi non orientati, il determinante tropicale del laplaciano ha una struttura particolarmente ricca legata alla matrice degli spanning tree (formula matriciale di Kirchhoff). Il risultato fondamentale è:

$$\det_{\text{trop}}(L) = \max\{\text{peso tropicale di un spanning tree di } G(X, \sim)\}$$

dove il peso tropicale di uno spanning tree T è la somma (reale) dei pesi degli archi di T . Nel caso non pesato (tutti gli archi di peso 1), il peso tropicale è uguale al numero di archi del spanning tree (cioè $n-1$ per ogni spanning tree), e $\det_{\text{trop}}(L) = n-1$ indipendentemente dalla struttura del grafo — un risultato degenerare che riflette la mancanza di pesi nell'algebra tropicale non pesata.

Per grafi pesati (dove il peso dell'arco (i,j) è la distanza ordinale $|i-j|$ nel dominio del predicato), il determinante tropicale del laplaciano pesato è:

$$\det_{\text{trop}}(L_w) = \max_T \{ \sum_{(i,j) \in T} |i-j| \}$$

il massimo spanning tree in peso degli archi — il «cammino di Prüfer massimale» nello spazio ordinale. Questo peso massimale è collegato alla bandwidth del grafo: $\det_{\text{trop}}(L_w) = \text{bw}_{\leq}(\sim) \cdot (n-1)$ per grafi con struttura uniforme.

Teorema XVI.1 (Dualità tropicale tra ε^* e il determinante tropicale). Per uno spazio di tolleranza $(X, \sim, \leq)$ con n elementi e grafo $G(X, \sim)$ connesso, il numero di risoluzione ε^* e il determinante tropicale del laplaciano pesato L_w soddisfano la relazione duale: $\varepsilon^*(\sim) \cdot \det_{\text{trop}}(L_w) = \max_T \{ (1/L_{\text{mono_min}}) \cdot \sum_{(i,j) \in T} |i-j| \}$, dove il massimo è sul massimo spanning tree T di $G_w(X, \sim)$. Per tolleranze uniformi di larghezza k su n elementi: $\varepsilon^*(\sim) \cdot \det_{\text{trop}}(L_w) = (k/(n-1)) \cdot k(n-1) = k^2$ — il quadrato della larghezza di tolleranza. Il prodotto $\varepsilon^* \cdot \det_{\text{trop}}(L_w)$ è quindi un invariante «misto» che cattura simultaneamente la struttura locale ($\det_{\text{trop}}$) e quella globale (ε^*).

XVI.3 La varietà tropicale del sorites e la fan di Gröbner

La fan di Gröbner di un ideale polinomiale $I \subset k[x_1, \dots, x_n]$ è la partizione del cono positivo $\mathbb{R}_{\geq 0}^n$ delle «direzioni di peso» in regioni su cui la base di Gröbner dell'ideale I rimane costante. La geometria tropicale stabilisce (Sturmfels, 2002) che la varietà tropicale di un ideale I è la fan di Gröbner duale dell'ideale: $V_{\text{trop}}(I) = \{w \in \mathbb{R}^n : \text{in}_w(I) \neq \{1\}\}$.

Per tradurre questa struttura nel contesto delle strutture di tolleranza, si definisce l'ideale di tolleranza $I(X, \sim) \subset \mathbb{R}[f_1, \dots, f_n]$ come l'ideale generato dai binomi di coerenza:

$$I(X, \sim) = \langle f_i - f_j : x_i \sim x_j \rangle \subset \mathbb{R}[f_1, \dots, f_n]$$

L'ideale $I(X, \sim)$ è un ideale lineare — generato da polinomi di grado 1 — e la sua varietà (il luogo degli zeri in $\mathbb{R}^n$) è il sottospazio lineare degli assegnamenti coerenti: $\text{Var}(I) = \{f \in \mathbb{R}^n : f(x_i) = f(x_j) \text{ per tutti } x_i \sim x_j\} = \bigoplus_{\text{componenti } C} \{1_C\}$, isomorfo a $\mathbb{R}^k$ (dove k è il numero di componenti connesse). Questo è l'analogo algebrico della Proposizione 2.

La varietà tropicale $V_{\text{trop}}(I(X, \sim))$ dell'ideale di tolleranza è la fan di Gröbner dell'ideale lineare, che per ideali lineari coincide con la fan normale del politopo delle soluzioni. Nel caso dello spazio di tolleranza, la fan normale è la fan dei coni di Weyl della decomposizione dell'ipercubo $[0,1]^n$ indotta dai vincoli di coerenza. Ogni cono massimale della fan corrisponde a un possibile ordinamento delle funzioni di membership compatibile con la struttura di tolleranza.

La fan di Gröbner dell'ideale di tolleranza ha una struttura combinatoria precisa: i suoi coni massimali sono in biiezione con gli ordinamenti lineari degli elementi di X compatibili con la chiusura transitiva della relazione di tolleranza. Per predicati unidimensionali (il caso principale di questo saggio), questo significa che i coni massimali corrispondono agli ordinamenti compatibili con l'ordinamento $\leq$ del dominio, e il numero di coni massimali è il numero di estensioni lineari del preordine $\sim^T \leq$ (la composizione della chiusura transitiva di $\sim$ con l'ordinamento del dominio).

Proposizione XVI.1 (Fan di Gröbner e numero di risoluzioni). La fan di Gröbner $F(X, \sim)$ dell'ideale di tolleranza $I(X, \sim)$ ha le seguenti proprietà: (i) il numero di coni massimali è il numero di estensioni lineari del preordine indotto da $\sim$; (ii) i vettori di peso w sulla frontiera tra due coni adiacenti corrispondono alle coppie tolleranti $x_i \sim x_j$ per cui il peso relativo $w_i - w_j$ cambia segno; (iii) la larghezza minima di un cono massimale (misurata rispetto alla norma euclidea) è proporzionale a $\varepsilon^*(\sim)$: $\text{largh_min}(F) = \varepsilon^*(\sim) \cdot \sqrt{n}$. Il numero di risoluzione è dunque inversamente proporzionale alla 'finezza' della fan di Gröbner: spazi di tolleranza più vaghi (ε^* piccolo) hanno fan più fini (coni più stretti).

XVI.4 La struttura delle basi di Gröbner tropicali e l'algoritmo di Buchberger tropicale

L'algoritmo di Buchberger (1965) per il calcolo di basi di Gröbner ha un analogo tropicale, sviluppato da Jensen, Markwig e Markwig (2008) per ideali tropicali. Per l'ideale di tolleranza $I(X, \sim)$, la base di Gröbner tropicale si calcola come segue.

Si considera il generatore $\xi_{ij} = f_i - f_j$ per ogni coppia tollerante $x_i \sim x_j$. L'S-polinomio di due generatori ξ_{ij} e ξ_{jk} è:

$$S(\xi_{ij}, \xi_{jk}) = f_i - f_j - (f_j - f_k) = f_i - 2f_j + f_k$$

che non è nell'ideale $I(X, \sim)$ a meno che esista un generatore che lo riduce a zero. Specificamente, la riduzione è possibile se esiste una catena $f_i \sim f_j \sim f_k$ con $f_i \sim f_k$ (transitività), nel qual caso $\xi_{ik} = f_i - f_k$ e $S(\xi_{ij}, \xi_{jk})$ si riduce a zero tramite la combinazione lineare $\xi_{ij} + \xi_{jk} - \xi_{ik} = (f_i - f_j) + (f_j - f_k) - (f_i - f_k) = 0$. Se invece $f_i \not\sim f_k$ (non-transitività della tolleranza!), l'S-polinomio non si riduce a zero e deve essere aggiunto alla base di Gröbner come generatore aggiuntivo. Questo è il punto cruciale: la non-transitività della relazione di tolleranza produce S-polinomi che non si riducono a zero — e quindi obbliga l'algoritmo di Buchberger ad aggiungere elementi alla base.

La base di Gröbner ridotta dell'ideale di tolleranza $I(X, \sim)$ è l'insieme:

$$GB(I(X, \sim)) = \{ f_i - f_j : x_i \text{ e } x_j \text{ sono nella stessa componente} \\ \text{connessa di } G(X, \sim) \}$$

Questa base è l'insieme di tutti i binomi che esprimono la coerenza tra ogni coppia di elementi nella stessa componente connessa — incluse le coppie che non sono direttamente tolleranti. Il numero di elementi nella base di Gröbner è $\sum_k C(|C_k|, 2) = \sum_k |C_k|(|C_k|-1)/2$, dove C_k sono le componenti connesse e $C(n,2)$ è il numero di coppie. Per un grafo connesso, il numero di elementi nella base è $C(n,2) = n(n-1)/2$ — il numero di coppie di vertici. La base di Gröbner è dunque molto più grande dell'insieme dei generatori originali (che sono solo $|E| \leq n(n-1)/2$), e la differenza $|GB| - |E|$ misura il «deficit di transitività» della relazione di tolleranza.

Teorema XVI.2 (Deficit di transitività e vaghezza). Per uno spazio di tolleranza $(X, \sim)$ con grafo $G(X, \sim)$ connesso e n elementi, il deficit di transitività è: $\Delta_{\text{transit}}(\sim) = |GB(I(X, \sim))| - |E(G(X, \sim))| = n(n-1)/2 - |E|$. Il deficit di transitività è zero se e solo se $G(X, \sim) = K_n$ (il grafo completo), cioè se $\sim$ è la relazione universale. Il numero di risoluzione ε^* è limitato superiormente dal deficit normalizzato: $\varepsilon^*(\sim) \leq 1 - 2|E|/(n(n-1)) = \Delta_{\text{transit}}(\sim) / C(n,2)$. La vaghezza (misurata da ε^*) è tanto maggiore quanto minore è la densità del grafo di tolleranza rispetto al grafo completo — cioè quanto maggiore è il deficit di transitività.

XVI.5 Poliedri Newton-Okounkov e i corpi di valutazione dei predicati vaghi

I corpi di Newton-Okounkov (Lazarsfeld e Mustață, 2009; Kaveh e Khovanskii, 2012) generalizzano i politopi di Newton dei polinomi a una classe molto più ampia di «corpi convessi associati ad algebre graduate». Per lo spazio dei predicati vaghi, il corpo di Newton-Okounkov è definito a partire dalla valutazione tropicale della funzione di membership.

Si considera la famiglia di spazi vettoriali $V_m = \{\text{polinomi in } f_1, \dots, f_n \text{ di grado } \leq m \text{ che si annullano sull'ideale } I(X, \sim)\} \subset \mathbb{R}[f_1, \dots, f_n]_m$. La valutazione di Gröbner $v: V_m \setminus \{0\} \rightarrow \mathbb{Z}^n$ (rispetto a un ordinamento monomiale) associa a ogni polinomio il suo termine guida (monomo con maggiore peso). Il corpo di Newton-Okounkov $\Delta(\mathcal{R}, v)$ associato all'algebra $\mathcal{R} = \bigoplus_m V_m$ è il limite (nel senso di Hausdorff) della sequenza di politopi:

$$\Delta_m = (1/m) \cdot \text{conv}\{ v(f) : f \in V_m \setminus \{0\} \} \rightarrow \Delta(\mathcal{R}, v) \text{ per } m \rightarrow \infty$$

Per l'ideale di tolleranza $I(X, \sim)$ con ordinamento monomiale lessicografico, il corpo di Newton-Okounkov è il politopo convesso in $\mathbb{R}^k$ (con k il numero di componenti connesse) i cui vertici corrispondono agli assegnamenti di valori 0 o 1 alle componenti. Questo politopo è l'ipercubo $[0,1]^k$ — esattamente lo spazio degli assegnamenti coerenti bivalenti, le cui 2^k soluzioni intere sono gli assegnamenti coerenti della Proposizione 1. Il corpo di Newton-Okounkov codifica dunque la struttura combinatoria degli assegnamenti coerenti in una forma geometrica.

Il volume del corpo di Newton-Okounkov è 1 (il volume dell'ipercubo unitario in dimensione k), indipendentemente dalla struttura interna delle componenti connesse. Questo riflette il fatto che il

numero di assegnamenti coerenti bivalenti è 2^k per qualunque struttura interna delle componenti. Tuttavia, il volume del corpo di Newton-Okounkov per assegnamenti ε -coerenti (il politopo ε -deformato) è:

$$\text{Vol}(\Delta(\varepsilon)) = (1 + \varepsilon \cdot L_{\text{mono_min}})^k / k! \approx (1 + 1)^k / k! = 2^k / k! \\ \text{per } \varepsilon = \varepsilon^* = 1/L_{\text{mono_min}}$$

che misura il volume dello spazio degli assegnamenti ε -coerenti con contrasto — la «densità» delle soluzioni ammissibili nel problema di risoluzione della vaghezza.

.

PARTE XVII. ANALISI NON-STANDARD E INFINITESIMALI SEMANTICI: LA VAGHEZZA NELL'IPERPIANO DI ROBINSON

XVII.1 I fondamenti dell'analisi non-standard e il principio di trasferimento

L'analisi non-standard, fondata da Abraham Robinson nel 1961 e sistematizzata nel testo *Nonstandard Analysis* (1966), costruisce un'estensione ${}^*\mathbb{R}$ dei reali $\mathbb{R}$ che contiene elementi infinitamente piccoli (infinitesimali) e infinitamente grandi (infiniti), dotati di tutte le proprietà dei reali ordinari nel senso del Principio di Trasferimento. Formalmente, ${}^*\mathbb{R}$ è un campo ordinato non archimedeo che estende $\mathbb{R}$: $\mathbb{R} \subset {}^*\mathbb{R}$, e ${}^*\mathbb{R}$ contiene elementi $\varepsilon > 0$ tali che $\varepsilon < r$ per ogni $r \in \mathbb{R}$, $r > 0$ (gli infinitesimali positivi) e elementi $\Omega > 0$ tali che $\Omega > r$ per ogni $r \in \mathbb{R}$ (gli infiniti positivi).

Il Principio di Trasferimento (Łoś, 1955; Robinson, 1966) afferma che ogni enunciato del primo ordine nel linguaggio dei campi ordinati è vero in $\mathbb{R}$ se e solo se è vero in ${}^*\mathbb{R}$. Questo garantisce che ${}^*\mathbb{R}$ si comporta esattamente come $\mathbb{R}$ per tutte le proprietà esprimibili nel primo ordine: i teoremi dell'algebra, dell'analisi reale, delle equazioni differenziali che si possono enunciare nel primo ordine si trasferiscono automaticamente a ${}^*\mathbb{R}$. La differenza con $\mathbb{R}$ emerge negli enunciati del secondo ordine (che quantificano su sottoinsiemi) e nelle proprietà di completezza (ogni insieme limitato di $\mathbb{R}$ ha sup, ma ${}^*\mathbb{R}$ ha elementi limitati senza sup in ${}^*\mathbb{R}$).

Per la teoria degli spazi di tolleranza, l'analisi non-standard fornisce un quadro in cui la «soglia di indistinguibilità» può essere modelata come un numero infinitesimale $\varepsilon \in {}^*\mathbb{R}$, $\varepsilon > 0$, $\varepsilon < r$ per ogni r reale positivo. Questo corrisponde alla situazione limite in cui la granularità del predicato diventa infinitamente fine — ma non zero. Nel caso continuo ($X = [0,1]$ con la tolleranza $|x-y| \leq \varepsilon$ per ε infinitesimale), il paradosso di Sorite diventa immediatamente paradossale: la catena di tolleranza da 0 a 1 richiede un numero iperfinito di passi $N \sim 1/\varepsilon \in {}^*\mathbb{N} \setminus \mathbb{N}$, e il numero di risoluzione $\varepsilon^* = 1/N \sim \varepsilon$ è infinitesimale ma positivo.

XVII.2 La costruzione degli spazi di tolleranza iperfiniti

Si consideri l'estensione iperfinita ${}^*X$ del dominio finito $X = \{1, \dots, n\}$: nello spirito dell'analisi non-standard, si prende un'ultrafiltro non principale U su $\mathbb{N}$ e si costruisce il prodotto ultrafiltrante ${}^*X = \prod_n X_n / U$, dove $X_n = \{1, \dots, n\}$ per ogni n . Il risultato è un insieme iperfinito ${}^*X$ che «contiene $\mathbb{N}$ copie di X » nel senso dell'ultrafiltro.

Più direttamente, si considera il caso continuo $X = [0, 1]$ con la tolleranza iperfinita: si fissa un infinitesimale $\varepsilon > 0$ in ${}^*\mathbb{R}$ e si definisce la relazione di tolleranza iperfinita $\sim_{\varepsilon}$ su ${}^*[0, 1]$ tramite:

$$x \sim_{\varepsilon} y \iff {}^*|x - y| \leq \varepsilon, \quad x, y \in {}^*[0, 1]$$

dove ${}^*|\cdot|$ è il valore assoluto esteso a ${}^*\mathbb{R}$ tramite il Principio di Trasferimento. Lo spazio di tolleranza $({}^*[0, 1], \sim_{\varepsilon})$ è uno spazio di tolleranza continuo su un campo non archimedeo. Il suo grafo di tolleranza $G({}^*[0, 1], \sim_{\varepsilon})$ è il grafo su ${}^*[0, 1]$ dove due punti sono adiacenti se distano al più ε .

Il paradosso di Sorite nel contesto non-standard prende la seguente forma: la catena di tolleranza da 0 a 1 ha lunghezza iperfinita $N = \lceil 1/\varepsilon \rceil \in {}^*\mathbb{N} \setminus \mathbb{N}$. Il numero di risoluzione $\varepsilon^*(\sim_{\varepsilon}) = 1/N \sim \varepsilon$ è infinitesimale. L'assegnamento ottimale $f^*: {}^*[0, 1] \rightarrow {}^*[0, 1]$ è la funzione identità $f^*(x) = x$ (la funzione di Łukasiewicz con valori in ${}^*[0, 1]$), che è ε -coerente con $\varepsilon =$ infinitesimale.

Teorema XVII.1 (Il paradosso di Sorite nell'analisi non-standard). Per lo spazio di tolleranza $({}^*[0, 1], \sim_{\varepsilon})$ con $\varepsilon \in {}^*\mathbb{R}$ infinitesimale positivo, vale: (i) $G({}^*[0, 1], \sim_{\varepsilon})$ è connesso (il paradosso di Sorite si applica); (ii) $\varepsilon^*(\sim_{\varepsilon}) = \varepsilon$ è infinitesimale ma non zero; (iii) l'assegnamento ottimale $f^*: {}^*[0, 1] \rightarrow {}^*[0, 1]$ è la funzione $f^*(x) = x$, che è ε -coerente con ε infinitesimale; (iv) la parte standard $\text{st}(f^*(x)) = \text{st}(x) \in [0, 1]$ dell'assegnamento ottimale è la funzione identità su $[0, 1]$ — esattamente la funzione di Łukasiewicz; (v) la parte standard $\text{st}(\varepsilon^*(\sim_{\varepsilon})) = 0$: nel senso standard, il numero di risoluzione 'collapsa a zero' nell'analisi non-standard, riflettendo il fatto che la vaghezza è infinitesimale ma genuinamente non-nulla.

XVII.3 Gli infinitesimali semantici e la gerarchia di vaghezza

L'analisi non-standard fornisce una gerarchia naturale di «gradi di vaghezza» corrispondente alla gerarchia degli ordini di infinitesimità. Due infinitesimali ε_1 e ε_2 si dicono dello stesso ordine se $\varepsilon_1/\varepsilon_2$ è un numero reale finito non nullo; ε_1 è di ordine superiore (più piccolo) di ε_2 se $\varepsilon_1/\varepsilon_2$ è infinitesimale. Questa gerarchia si traduce in una gerarchia di predicati vaghi:

Un predicato P_1 con numero di risoluzione $\varepsilon^*(P_1) = \varepsilon_1$ e un predicato P_2 con $\varepsilon^*(P_2) = \varepsilon_2$ sono dello stesso «grado di vaghezza» se $\varepsilon_1/\varepsilon_2$ è finito; P_1 è «più vago» di P_2 se $\varepsilon^*(P_1)/\varepsilon^*(P_2)$ è infinitesimale. Per i predicati standard (con ε^* reale positivo), tutti sono allo stesso «grado» nell'analisi non-

standard. Il fenomeno interessante emerge per predicati con ε^* reale infinitesimale — predicati la cui vaghezza è «al limite del percettibile» ma non eliminabile.

Si definisce la gerarchia di Hrbáček-Nelson dei predicati vaghi (ispirata alla gerarchia delle «galassie» nell'analisi non-standard): un predicato P è n -standard ($n \in \mathbb{N}$) se il suo numero di risoluzione $\varepsilon^*(P)$ è di ordine ε^n (con ε un infinitesimale fissato). La gerarchia di vaghezza è:

$$\begin{aligned} 0\text{-standard (standard)}: & \varepsilon^*(P) \in \mathbb{R}^+ \\ 1\text{-standard (quasi-nitido)}: & \varepsilon^*(P) \sim \varepsilon \quad (\text{infinitesimale del} \\ & \text{prim'ordine}) \\ 2\text{-standard (iper-vago)}: & \varepsilon^*(P) \sim \varepsilon^2 \quad (\text{infinitesimale del} \\ & \text{second'ordine}) \\ k\text{-standard}: & \varepsilon^*(P) \sim \varepsilon^k \quad (\text{infinitesimale del } k\text{-esimo ordine}) \end{aligned}$$

Questa gerarchia ha un'interpretazione percettiva concreta: un predicato k -standard è uno il cui confine semantico è percepibile solo dopo k iterazioni di «zoom psicofisico» — k applicazioni della JND. I predicati più comuni del linguaggio naturale (alto, basso, caldo, freddo, rosso, verde) sono 0-standard: la loro vaghezza è percepibile direttamente. I predicati come «altissimo» o «caldissimo» (con superlativo) tendono a essere 1-standard: il loro confine è più fine ma ancora percepibile. I predicati tecnici altamente specializzati (come soglie regolamentari in medicina o diritto) tendono a essere quasi-0-standard: la loro vaghezza residua è quasi-nulla.

XVII.4 Il teorema di compattezza per spazi di tolleranza e la vaghezza globale

Il teorema di compattezza della logica del primo ordine (Gödel, 1930; Malcev, 1936) afferma che un insieme di formule Γ del primo ordine è soddisfacibile se e solo se ogni suo sottoinsieme finito è soddisfacibile. Nell'analisi non-standard, questo teorema ha la seguente riformulazione (Robinson, 1966): un insieme di enunciati del primo ordine è soddisfacibile nel campo ${}^*\mathbb{R}$ se e solo se ogni suo sottoinsieme finito è soddisfacibile in $\mathbb{R}$.

Applicato alla teoria degli spazi di tolleranza, il teorema di compattezza dà il seguente risultato fondamentale:

Teorema XVII.2 (Compattezza per la vaghezza — Teorema di Robinson semantico).

Sia P un predicato vago con struttura di tolleranza $\sim$ su un dominio infinito X (ad esempio $X = \mathbb{N}$ o $X = [0,1]$). P ammette un assegnamento 0-coerente non costante su X (cioè il paradosso di Sorite non si applica) se e solo se ogni restrizione di P a un sottodominio finito $X_{\text{fin}} \subset X$ ammette un assegnamento 0-coerente non costante. In altri termini: il paradosso di Sorite per un predicato su un dominio infinito equivale al paradosso di Sorite su ogni suo sottodominio finito. La vaghezza è un fenomeno 'finitariamente compatto': emerge già su insiemi finiti e non richiede il dominio infinito.

Dimostrazione. Sia Γ l'insieme di enunciati del primo ordine che descrivono lo spazio di tolleranza $(X, \sim)$ e il fatto che esiste un assegnamento 0-coerente non costante. Per il teorema di compattezza, Γ è soddisfacibile se e solo se ogni suo sottoinsieme finito $\Gamma_{\text{fin}} \subset \Gamma$ è soddisfacibile. Ma ogni sottoinsieme finito di Γ descrive soltanto un numero finito di elementi di X e un numero finito di relazioni di tolleranza tra essi — quindi descrive un sottodominio finito $X_{\text{fin}} \subset X$. L'equivalenza segue direttamente. ■

Il Teorema XVII.2 ha un corollario importante per la linguistica computazionale: per analizzare la vaghezza di un predicato è sufficiente studiarne le proprietà su campioni finiti del dominio. Non c'è nessuna proprietà di vaghezza «globale» che emerge soltanto nel limite continuo e non è rilevabile su campioni finiti — la vaghezza è un fenomeno intrinsecamente finito. Questo giustifica retroattivamente l'approccio dell'intero saggio: lavorare su domini finiti X è senza perdita di generalità per le proprietà di vaghezza del primo ordine.

XVII.5 Gli iperreali nella modellazione della JND continua e il paradosso di Sorite iper-standard

La psicofisica classica descrive la JND (just noticeable difference) come una quantità reale positiva — il minimo incremento percepibile di uno stimolo. Ma experimentalmente, la JND dipende dall'intensità dello stimolo (legge di Weber: $JND/I = \text{costante } k$) e può diventare arbitrariamente piccola per stimoli di intensità elevata. Nel limite $I \rightarrow \infty$, la $JND \rightarrow 0$ in senso reale, ma la distinzione tra « $JND = 0$ » e «JND infinitesimale» è esattamente la distinzione catturata dall'analisi non-standard.

Si modella la JND come un infinitesimale $\eta \in {}^*\mathbb{R}$ nel senso di Robinson: $\eta > 0$ ma $\eta < r$ per ogni $r \in \mathbb{R}$, $r > 0$. Lo spazio di tolleranza iperfinito $({}^*I, \sim_\eta)$ dove ${}^*I = [I_{\text{min}}, I_{\text{max}}]$ e $x \sim_\eta y \Leftrightarrow |x - y| \leq \eta$ ha la seguente struttura: la lunghezza del cammino monotono minimo da I_{min} a I_{max} è $N = [(I_{\text{max}} - I_{\text{min}})/\eta] \in {}^*\mathbb{N} \setminus \mathbb{N}$ (un numero iperfinito), il numero di risoluzione è $\varepsilon^*(\sim_\eta) = 1/N \sim \eta/(I_{\text{max}} - I_{\text{min}})$ (infinitesimale), e il paradosso di Sorite è iper-standard: la catena di tolleranza ha lunghezza iperfinita. Il predicato ottimale $f^*(x) = (x - I_{\text{min}})/(I_{\text{max}} - I_{\text{min}})$ è ε -coerente con $\varepsilon \sim \eta/(I_{\text{max}} - I_{\text{min}})$ infinitesimale.

Il punto interessante è la relazione tra la JND continua (η infinitesimale in ${}^*\mathbb{R}$) e la JND discreta (h reale positivo in $\mathbb{R}$): la teoria discreta si recupera come parte standard della teoria iper-standard. La funzione di membership ottimale discreta f^*_{disc} (con dominio discreto $X_{\text{disc}} = \{I_{\text{min}}, I_{\text{min}}+h, I_{\text{min}}+2h, \dots\}$ e JND discreta h) converge, nell'analisi non-standard, alla funzione continua $f^*(x)$ per $h \rightarrow 0$ nel senso ultrafiltrante. Questo mostra che i risultati della teoria discreta non sono artefatti della discretizzazione ma riflettono la struttura continua sottostante.

La definizione di Cauchy della continuità si traduce direttamente nel contesto semantico: una funzione di membership $f: [0,1] \rightarrow [0,1]$ è continua in senso standard se e solo se per ogni coppia (x, y) con $x \sim_\eta y$ (infinitesimalmente vicini), si ha $f(x) \sim f(y)$ (infinitesimalmente vicini nel codominio). Questo è esattamente la condizione di 0-coerenza per la tolleranza iperfinita $\sim_\eta$: la

continuità ordinaria è la coerenza infinitesimale. Il paradosso di Sorite nel dominio continuo è dunque equivalente alla non-esistenza di funzioni continue non-costanti da $[0,1]$ con $f(0) = 0$ e $f(1) = 1$ che siano 0-coerenti rispetto a $\sim_\eta$ — ma tutte le funzioni continue non-costanti con valori agli estremi 0 e 1 sono ε -coerenti con $\varepsilon = \eta$ (infinitesimale), e il paradosso di Sorite nel senso iper-standard ha costo infinitesimale.

.

PARTE XVIII. TEORIA ERGODICA DELLA VAGHEZZA: MEDIE TEMPORALI, ENTROPIA DI KOLMOGOROV-SINAI E SISTEMI DINAMICI SEMANTICI

XVIII.1 Lo spazio delle misure su $\mathcal{F}_{\varepsilon^*}$ e le trasformazioni semantiche

La teoria ergodica studia il comportamento a lungo termine di sistemi dinamici — trasformazioni $T: X \rightarrow X$ su uno spazio con una misura μ invariante ($T^*\mu = \mu$) — tramite le proprietà statistiche delle orbite $\{T^n x\}_{n \geq 0}$. Il teorema di Birkhoff (1931) afferma che per quasi ogni punto $x \in X$ (rispetto a μ), la media temporale $(1/N) \sum_{n=0}^{N-1} f(T^n x)$ converge alla media spaziale $\int f d\mu$ quando il sistema è ergodico (l'unica misura T -invariante è μ). Applicata agli spazi di predicati vaghi, la teoria ergodica studia come le classificazioni di un parlante (o di una comunità linguistica) evolvono nel tempo.

Si modella l'evoluzione temporale di un predicato vago come una trasformazione T sul poliedro degli assegnamenti $\mathcal{F}_{\varepsilon^*}$. La trasformazione T rappresenta un «passo semantico»: l'aggiornamento del predicato dopo una nuova classificazione, un cambiamento di contesto, o una perturbazione del consenso linguistico. Le trasformazioni naturali da considerare sono:

(a) La trasformazione di rimescolamento semantico $T_{\text{rand}}: f \mapsto f + \eta$, dove η è un vettore di rumore gaussiano ε^* -contenuto ($|\eta_i| \leq \varepsilon^*$ per ogni i). Questa modella il rumore di classificazione: i parlanti variano le loro classificazioni in modo casuale entro i limiti della coerenza.

- (b) La trasformazione di proiezione $T_{\text{proj}}: f \mapsto \Pi_{\{\mathcal{F}_{\varepsilon^*}\}}(f + \eta)$, dove Π è la proiezione euclidea sul poliedro $\mathcal{F}_{\varepsilon^*}$. Questa modella la «correzione di coerenza»: dopo una perturbazione, il parlante corregge il proprio assegnamento al predicato più vicino tra quelli coerenti.
- (c) La trasformazione di consenso $T_{\text{cons}}: f \mapsto (1-\alpha)f + \alpha \cdot \bar{f}$, dove $\bar{f}$ è la media degli assegnamenti della comunità linguistica e $\alpha \in (0,1)$ è il tasso di convergenza verso il consenso. Questa modella la pressione sociale verso l'uniformità linguistica.

Definizione XVIII.1 (Sistema dinamico semantico). Un sistema dinamico semantico è una tripla $(\mathcal{F}_{\varepsilon^*}, T, \mu)$, dove $\mathcal{F}_{\varepsilon^*} \subset [0,1]^n$ è il poliedro degli assegnamenti ε^* -coerenti monotoni con contrasto, $T: \mathcal{F}_{\varepsilon^*} \rightarrow \mathcal{F}_{\varepsilon^*}$ è una trasformazione misurabile che preserva la struttura del poliedro, e μ è una misura di probabilità T -invariante su $\mathcal{F}_{\varepsilon^*}$ ($T^*\mu = \mu$).

XVIII.2 Il teorema di Birkhoff e la media temporale degli assegnamenti vaghi

Per il sistema dinamico semantico $(\mathcal{F}_{\varepsilon^*}, T_{\text{cons}}, \mu_{\text{inv}})$, dove μ_{inv} è la misura di Liouville (proporzionale alla misura di Lebesgue su $\mathcal{F}_{\varepsilon^*}$), il teorema di Birkhoff dà:

Teorema XVIII.1 (Birkhoff semantico). Sia $(\mathcal{F}_{\varepsilon^*}, T_{\text{cons}}, \mu_{\text{Liouv}})$ il sistema dinamico semantico con la trasformazione di consenso T_{cons} con tasso $\alpha \in (0,1)$. Per μ_{Liouv} -quasi ogni assegnamento iniziale $f_0 \in \mathcal{F}_{\varepsilon^*}$, la media temporale (il «predicato medio» di lungo periodo) converge: $\bar{f}_{\infty} = \lim_{N \rightarrow \infty} (1/N) \sum_{n=0}^{N-1} T_{\text{cons}}^n(f_0) = \int_{\mathcal{F}_{\varepsilon^*}} f \, d\mu_{\text{inv}} = f^*_{\text{baricentro}}$. Il predicato di lungo periodo è il baricentro del poliedro $\mathcal{F}_{\varepsilon^*}$ rispetto alla misura di Liouville, che coincide con l'assegnamento ottimale $f^* = 1/L_{\text{mono_min}} \cdot (j(x_i))$ (la funzione gradino a passi $1/L$ del Capitolo 5) per sistemi ergodici. Per sistemi non ergodici, la media converge a una media che dipende dal bacino di attrazione dell'orbita.

La dimostrazione richiede di verificare che T_{cons} è una contrazione lipschitziana su $\mathcal{F}_{\varepsilon^*}$ (con costante di Lipschitz $1-\alpha < 1$) e che la misura di Liouville è T_{cons} -invariante nel limite $\alpha \rightarrow 0$. Per α fissato, T_{cons} ha un unico punto fisso $f^* = \bar{f}_{\text{comunità}}$ (il predicato di consenso della comunità) e tutte le orbite convergono esponenzialmente verso f^* con tasso $(1-\alpha)^n$. ■

XVIII.3 L'entropia di Kolmogorov-Sinai e la complessità della vaghezza

L'entropia di Kolmogorov-Sinai (KS-entropia) di un sistema dinamico è l'analogo ergodico dell'entropia di Shannon: misura la quantità media di informazione generata per passo dalla traiettoria del sistema. Per un sistema dinamico (X, T, μ) con partizione $P = \{P_1, \dots, P_k\}$ di X , l'entropia rispetto a P è:

$$h(T, P) = \lim_{n \rightarrow \infty} (1/n) H(P \vee T^{-1}P \vee \dots \vee T^{-(n-1)}P)$$

dove $P \vee Q$ è il raffinamento comune delle partizioni P e Q e $H(Q) = -\sum_i \mu(Q_i) \log \mu(Q_i)$ è l'entropia di Shannon della partizione Q . L'entropia KS è il supremo di $h(T, P)$ su tutte le partizioni finitemente generate: $h_{KS}(T) = \sup_P h(T, P)$.

Per il sistema dinamico semantico $(\mathcal{F}_{\varepsilon^*}, T, \mu_{Liouv})$, la partizione naturale da considerare è la partizione in «zone di verità»: la partizione P_{ε^*} di $\mathcal{F}_{\varepsilon^*}$ in tre regioni — la zona chiaramente vera ($\{f \in \mathcal{F}_{\varepsilon^*} : f(x_n) > 1 - \varepsilon^*\}$), la zona di confine ($\{f : \varepsilon^* \leq f(x_i) \leq 1 - \varepsilon^* \text{ per qualche } i\}$) e la zona chiaramente falsa ($\{f : f(x_1) < \varepsilon^*\}$). L'entropia rispetto a questa partizione misura la «velocità con cui le classificazioni cambiano zona di verità».

Teorema XVIII.2 (Entropia KS e numero di risoluzione). Per il sistema dinamico semantico con trasformazione di rimescolamento T_{rand} e misura di Liouville μ_{Liouv} , l'entropia di Kolmogorov-Sinai rispetto alla partizione P_{ε^*} soddisfa: $h_{KS}(T_{rand}, P_{\varepsilon^*}) = \log(1/\varepsilon^*(\sim)) + C$, dove C è una costante che dipende dal volume del poliedro $\mathcal{F}_{\varepsilon^*}$ ma non dalla struttura della tolleranza. L'entropia KS è tanto maggiore quanto più piccolo è $\varepsilon^*(\sim)$: predicati più vaghi generano più informazione per passo — le classificazioni cambiano più frequentemente tra le zone di verità, e la traiettoria del sistema è più «caotica» nel senso dell'entropia.

Questo risultato stabilisce una connessione tra due misure di complessità della vaghezza: il numero di risoluzione ε^* (misura topologica, Capitolo 5) e l'entropia KS (misura ergodica). I predicati più vaghi sono sia più difficili da «risolvere» (ε^* piccolo) sia più caotici nella loro dinamica temporale (h_{KS} grande): la vaghezza semantica si manifesta come complessità sia spaziale (topologica) sia temporale (ergodica).

XVIII.4 Il teorema di Poincaré e il ritorno semantico

Il teorema di ricorrenza di Poincaré (1890) afferma che per ogni sistema dinamico (X, T, μ) con $\mu(X) < \infty$ e μ T -invariante, quasi ogni punto di X ritorna arbitrariamente vicino alla sua posizione iniziale dopo un numero finito di iterazioni. Applicato al sistema semantico, il teorema di Poincaré afferma:

Teorema XVIII.3 (Ricorrenza semantica di Poincaré). Per il sistema dinamico semantico $(\mathcal{F}_{\varepsilon^*}, T, \mu_{Liouv})$ con $\mu_{Liouv}(\mathcal{F}_{\varepsilon^*}) < \infty$ e T misura-preservante, per μ_{Liouv} -quasi ogni assegnamento iniziale $f_0 \in \mathcal{F}_{\varepsilon^*}$ e ogni intorno aperto $U \ni f_0$, esiste un tempo $n \in \mathbb{N}$ tale che $T^n(f_0) \in U$. In termini semantici: quasi ogni predicato «torna» arbitrariamente vicino al suo stato iniziale dopo un numero finito di aggiornamenti semantici. Il tempo di Poincaré $\tau(f_0, U) = \min\{n \geq 1 : T^n(f_0) \in U\}$ soddisfa: $E_{\mu}[\tau] = 1/\mu(U)$ (teorema di Kac). Per $U = B_{\varepsilon^*}(f_0)$ (la palla di raggio ε^* attorno a f_0), il tempo di ritorno medio è: $E[\tau] =$

$1/\mu(B_{\{\varepsilon^*\}}(f_0)) \sim (1/\varepsilon^*)^n / \text{Vol}(\mathcal{F}_{\{\varepsilon^*\}})$: il tempo medio tra due «stati quasi-identici» cresce come la potenza n -esima di $1/\varepsilon^*$, dove $n = \dim \mathcal{F}_{\{\varepsilon^*\}} = L_{\text{mono_min}} - 1$. Il ritorno semantico è tanto più lento quanto più fine è la zona di confine.

XVIII.5 Il teorema ergodico spettrale e la decomposizione di Koopman

L'operatore di Koopman $U_T: L^2(\mathcal{F}_{\{\varepsilon^*\}}, \mu) \rightarrow L^2(\mathcal{F}_{\{\varepsilon^*\}}, \mu)$ associato alla trasformazione T è l'isometria di pullback $U_T f = f \circ T$. La teoria spettrale dell'operatore di Koopman fornisce informazioni complete sul comportamento a lungo termine del sistema dinamico: le funzioni proprie di U_T corrispondono alle componenti «periodiche» delle traiettorie, e la decomposizione spettrale di U_T nel sottospazio delle medie nulle fornisce le correlazioni temporali del sistema.

Per il sistema dinamico semantico con $T = T_{\text{cons}}(\alpha)$, l'operatore di Koopman ha struttura spettrale:

$$\text{Spec}(U_{\{T_{\text{cons}}\}}) \subseteq \{z \in \mathbb{C} : |z| = 1\} \cup \{(1-\alpha)^k \lambda_k : \lambda_k \text{ autovalori di } L\}$$

dove $\{\lambda_k\}$ sono gli autovalori del laplaciano L del grafo di tolleranza. Il raggio spettrale di $U_{\{T_{\text{cons}}\}}$ nel sottospazio ortogonale alle costanti è:

$$r_{\text{spec}}(U_{\{T_{\text{cons}}\}}|_{\perp \text{const}}) = (1-\alpha)^{1/2} \cdot e^{\{-\lambda_1 / 2\}}$$

che è strettamente minore di 1 per $\alpha > 0$ e $\lambda_1 > 0$. Il sistema è dunque «mixing» (mescolante): le correlazioni temporali decadono esponenzialmente con tasso $-\log((1-\alpha)) - \lambda_1/2$, e il sistema «dimentica» il suo stato iniziale in tempo $\tau_{\text{mix}} \sim 1/(\alpha + \lambda_1/2)$. Per predicati più vaghi ($\lambda_1 \sim \varepsilon^{*2}$ piccolo) e con basso tasso di aggiornamento $\alpha \rightarrow 0$, il tempo di dimenticanza $\tau_{\text{mix}} \sim 1/\varepsilon^{*2}$ è dell'ordine quadratico inverso del numero di risoluzione.

Il teorema spettrale ergodico afferma che la convergenza delle medie temporali alla media spaziale avviene in norma L^2 con tasso r_{spec} : per ogni $f \in L^2(\mathcal{F}_{\{\varepsilon^*\}}, \mu)$, $\|(1/N) \sum_{n=0}^{N-1} U_T^n f - \int f d\mu\|_2 \leq r_{\text{spec}}^N \cdot \|f\|_2$. Questo è il Teorema di von Neumann (1932) nella versione ergodica: le classificazioni di lungo periodo convergono al consenso con tasso esponenziale governato dal gap spettrale λ_1 del laplaciano del grafo di tolleranza.

XVIII.6 Le misure SRB e la statistica degli attori linguistici

Le misure di Sinai-Ruelle-Bowen (misure SRB, Sinai 1972, Ruelle 1976, Bowen 1975) sono le misure invarianti fisicamente rilevanti per i sistemi dinamici caotici: sono le misure che descrivono la statistica delle traiettorie partenti da un insieme di condizioni iniziali di misura positiva di Lebesgue (non necessariamente di misura positiva per μ_{SRB} stessa). Per il sistema semantico, le misure SRB descrivono la distribuzione dei predicati osservati dalla prospettiva di «un parlante tipico».

Per il sistema dinamico semantico con trasformazione di rimescolamento T_{rand} (con rumore gaussiano di varianza $\sigma^2 = \varepsilon^{*2}$), la misura SRB è la misura di Gibbs con hamiltoniana $H(f) = \|Lf\|_2^2 / (2\sigma^2)$ (l'energia di Dirichlet normalizzata):

$$d\mu_{\text{SRB}}(f) = \frac{1}{Z_{\text{SRB}}} \cdot \exp(-H(f)) \cdot df = \frac{1}{Z_{\text{SRB}}} \cdot \exp(-\|Lf\|_2^2 / (2\varepsilon^{*2})) \cdot df$$

La misura SRB concentra la massa attorno ai predicati con bassa energia di Dirichlet — i predicati più «coerenti» nel senso del laplaciano. Il punto di massa maggiore è il predicato f^* che minimizza l'energia di Dirichlet sotto il vincolo di contrasto: questo è esattamente l'assegnamento ottimale del Capitolo 5, il che conferma che la misura SRB del sistema semantico è concentrata attorno al predicato ottimale.

La varianza della distribuzione dei valori di f attorno a f^* secondo la misura SRB è:

$$\text{Var}_{\text{SRB}}(f(x_i)) = \varepsilon^{*2} \cdot (L^{-1})_{ii} = \varepsilon^{*2} / \lambda_{\text{eff}}(x_i)$$

dove $\lambda_{\text{eff}}(x_i)$ è l'«autovalore effettivo» del laplaciano all'elemento x_i , dato dal reciproco della diagonale della matrice di Green $G = L^{-1}$ (la pseudo-inversa del laplaciano, definita sul sottospazio ortogonale alle costanti). Gli elementi con $\lambda_{\text{eff}}(x_i)$ piccolo (nella zona di confine, dove il laplaciano è «quasi degenere») hanno grande varianza di classificazione: sono gli elementi più difficili da classificare, quelli per cui i parlanti variano di più. Gli elementi agli estremi del dominio hanno λ_{eff} grande e piccola varianza: sono classificati in modo quasi-uniforme da tutti i parlanti.

PARTE XIX. TEORIA DEI NODI, INVARIANTI SKEIN E TOPOLOGIA DEI CICLI SEMANTICI NEI PREDICATI VAGHI

XIX.1 I cicli nel grafo di tolleranza come nodi e link

La teoria dei nodi (knot theory) studia le classi di equivalenza di curve chiuse semplici nello spazio tridimensionale $\mathbb{R}^3$ rispetto alle deformazioni continue (isotopie ambientali). Un nodo è una curva chiusa semplice non intersecante; un link è un'unione finita di nodi. L'invariante fondamentale che distingue i nodi è il polinomio di Alexander (1928), riformulato da Jones (1984) nel polinomio di Jones, e ulteriormente generalizzato al polinomio HOMFLY-PT (Freyd-Yetter-Hoste-Lickorish-Millett-Oceanu, 1985; Przytycki-Traczyk, 1987).

Il collegamento con la teoria degli spazi di tolleranza è meno diretto di quello con l'omologia simpliciale o la teoria spettrale, ma è matematicamente genuino: i cicli indipendenti del grafo di tolleranza $G(X, \sim)$ (che generano $H_1(S(X, \sim); \mathbb{Z})$, le ostruzioni omologiche del Capitolo 6) possono essere naturalmente immersi in $\mathbb{R}^3$ e interpretati come nodi o link. La topologia dei nodi fornisce invarianti aggiuntivi che distinguono cicli semantici non distinguibili dall'omologia (che è il caso, dato che tutti i cicli immersi nel piano $\mathbb{R}^2$ come semplici curve chiuse sono isotopi al cerchio standard).

Si fissa un'immersione del grafo di tolleranza $G(X, \sim)$ in $\mathbb{R}^3$: ogni vertice $x_i \in X$ è mandato in un punto $p(x_i) \in \mathbb{R}^3$, e ogni arco (x_i, x_j) è rappresentato da una curva liscia che connette $p(x_i)$ a $p(x_j)$. Per un grafo di tolleranza planare (che ammette un'immersione in $\mathbb{R}^2$ senza incroci), l'immersione in $\mathbb{R}^3$ è banale: i cicli sono tutti isotopi al cerchio standard e la teoria dei nodi non aggiunge informazione. Per grafi di tolleranza non-planari (ad esempio K_5 — il grafo completo su 5 vertici — o $K_{3,3}$ — il grafo bipartito completo — che per il teorema di Kuratowski sono i «mattoni» dei grafi non-planari), l'immersione in $\mathbb{R}^3$ produce strutture nodali non banali.

Definizione XIX.1 (Nodo semantico di un ciclo di tolleranza). Sia $C = (x_{i1}, x_{i2}, \dots, x_{ik}, x_{i1})$ un ciclo del grafo $G(X, \sim)$ (una sequenza ciclica di elementi con $x_{ij} \sim x_{i_{j+1}}$ per ogni j e $x_{ik} \sim x_{i1}$). Fissata un'immersione del grafo in $\mathbb{R}^3$, il nodo semantico del ciclo C è la curva chiusa $\gamma_C: S^1 \rightarrow \mathbb{R}^3$ ottenuta concatenando le curve che rappresentano gli archi (x_{i1}, x_{i2}) , (x_{i2}, x_{i3}) , ..., (x_{ik}, x_{i1}) . La classe di isotopia di γ_C in $\mathbb{R}^3$ è un invariante topologico del ciclo C (rispetto a deformazioni continue dell'immersione che non cambiano la struttura di incroci degli archi). Il polinomio di Jones $J(\gamma_C; q)$ e il polinomio HOMFLY-PT $P(\gamma_C; v, z)$ del nodo semantico γ_C sono invarianti topologici del ciclo C .

XIX.2 Il polinomio di Tutte come invariante combinatoriale degli spazi di tolleranza

Prima di addentrarsi nella teoria dei nodi propria, è opportuno presentare il polinomio di Tutte $T_G(x, y)$ del grafo di tolleranza $G(X, \sim)$, che è un invariante universale dei grafi rispetto ai minori e ha una relazione fondamentale con il polinomio di Jones dei nodi. Il polinomio di Tutte è definito ricorsivamente:

$$T_G(x, y) = T_{G \setminus e}(x, y) + T_{G/e}(x, y) \quad \text{se } e \text{ non è un ponte né un cappio}$$

$$T_G(x, y) = x \cdot T_{G \setminus e}(x, y) \quad \text{se } e \text{ è un ponte (arco la cui rimozione disconnette } G)$$

$$T_G(x, y) = y \cdot T_{G/e}(x, y) \quad \text{se } e \text{ è un cappio}$$

$$T_{K_n}(x, y) = x^{n-1} \quad (\text{grafo completo } K_n \text{ con } n-1 \text{ ponti})$$

dove $G \setminus e$ è il grafo ottenuto rimuovendo l'arco e e G/e è il grafo ottenuto contraendo l'arco e . Il polinomio di Tutte è un raffinamento del polinomio cromatico $P_G(k)$ e del polinomio di flusso

$F_G(k): P_G(k) = (-1)^{|V|-c(G)} k^{c(G)} T_G(1-k, 0)$ e $F_G(k) = (-1)^{|E|-|V|+c(G)} T_G(0, 1-k)$, dove $c(G)$ è il numero di componenti connesse.

Per lo spazio di tolleranza $(X, \sim)$ con grafo $G(X, \sim)$, il polinomio di Tutte $T_{\{G(X, \sim)\}}(x, y)$ è un invariante dello spazio di tolleranza. Le sue specializzazioni danno gli invarianti già incontrati nel saggio:

$T_G(1, 1)$ = numero di spanning tree di G (formula di Kirchhoff)

$T_G(2, 1)$ = numero di foreste di G (sottoinsiemi di archi aciclici)

$T_G(1, 2)$ = numero di archi-orientazioni acicliche di G (teorema di Stanley)

$T_G(2, 0)$ = numero di orientazioni di G senza cicli diretti (DAG con lo stesso grafo non orientato)

Quest'ultima specializzazione — $T_G(2, 0)$ — è particolarmente interessante per la teoria degli spazi di tolleranza: il numero di DAG senza cicli con il medesimo grafo sottostante $G(X, \sim)$ è il numero di ordinamenti lineari compatibili con la struttura del grafo, cioè il numero di estensioni lineari del preordine indotto da $\sim$. Per predicati unidimensionali (con X totalmente ordinato), questo numero è 1: c'è un unico DAG monotono compatibile con l'ordinamento, che è $G^*(X, \sim)$ stesso.

Teorema XIX.1 (Polinomio di Tutte e struttura della vaghezza). Per uno spazio di tolleranza $(X, \sim)$ con grafo G connesso e n elementi, il polinomio di Tutte $T_G(x, y)$ soddisfa: (i) $T_G(1, 1)$ = (numero di spanning tree) ≥ 1 ; (ii) $T_G(0, 2) = n^{n-2}$ per il grafo completo K_n (formula di Cayley); (iii) il coefficiente dominante di $T_G(x, y)$ rispetto a x^{n-1} è 1 per qualunque grafo connesso su n vertici; (iv) per grafi di tolleranza planari, $T_G(-1, -1) = (-1)^{|V|+1} \cdot 4^{\mu(G)/2} \cdot (\text{numero di 4-colorazioni di } G) / (\text{fattore di normalizzazione})$ — collegando il polinomio di Tutte con il teorema dei quattro colori (Appel-Haken, 1976); (v) la derivata parziale $\partial T_G / \partial x|_{(1,1)} = (\text{numero di spanning tree contenenti l'arco critico più critico})$ — che è legato alla probabilità di percolazione del grafo di tolleranza alla soglia di disconnessione.

XIX.3 La relazione di Kauffman e il polinomio di Jones dei cicli semantici

Il bracket di Kauffman $\langle D \rangle$ (Kauffman, 1987) di un diagramma di nodo D è il polinomio di Laurent in una variabile A definito dalla relazione di skein:

$$\langle \setminus \rangle = A \cdot \langle \rangle \langle \rangle + A^{-1} \cdot \langle \subset \supset \rangle \quad (\text{relazione di skein del bracket})$$

$$\langle D \sqcup \circ \rangle = (-A^2 - A^{-2}) \cdot \langle D \rangle \quad (\text{aggiunta di un componente non annodato})$$

$$\langle \circ \rangle = 1 \quad (\text{normalizzazione})$$

Il polinomio di Jones $V(q)$ si ottiene dal bracket di Kauffman mediante la normalizzazione che rende l'invariante rispetto alle mosse di Reidemeister di tipo I (le mosse di «arricciatura» che

modificano il writhe): $V_K(q) = (-A)^{-3w(D)} \cdot \langle D \rangle$, dove $w(D)$ è il writhe del diagramma D (la somma algebrica dei segni degli incroci) e $q = A^{-4}$.

Per i nodi semantici γ_C associati ai cicli del grafo di tolleranza, il calcolo del bracket di Kauffman richiede di fissare un diagramma dell'immersione di γ_C in $\mathbb{R}^3$ — cioè una proiezione di γ_C nel piano $\mathbb{R}^2$ con incroci. Per grafi di tolleranza planari, tutti i cicli si proiettano senza incroci e il bracket di Kauffman di ogni ciclo semantico è:

$$\langle \gamma_C \rangle_{\text{piano}} = (-A^2 - A^{-2})^{|C|-1} \quad \text{per cicli di lunghezza } |C| \text{ sul piano}$$

che corrisponde al polinomio del cerchio standard (ogni ciclo planare è isotopo al cerchio). Per grafi di tolleranza non-planari, i cicli semantici producono nodi non banali il cui bracket di Kauffman dipende dalla struttura degli incroci dell'immersione in $\mathbb{R}^3$.

Proposizione XIX.1 (Bracket di Kauffman per il ciclo fondamentale del sorite). Sia C_n il ciclo fondamentale del sorite su n elementi (il ciclo $(x_1, x_2, \dots, x_n, x_1)$ nel grafo $C_n = G(\{1, \dots, n\}, \sim)$ con $x_i \sim x_{i+1} \pmod{n}$). L'immersione canonica di C_n in $\mathbb{R}^2$ senza incroci dà il bracket di Kauffman: $\langle \gamma_{C_n} \rangle = (-A^2 - A^{-2})^{n-1}$, con polinomio di Jones $V(q) = (-q^{3/4})^{-3 \cdot 0} \cdot (-q^{1/2} - q^{-1/2})^{n-1} = (-1)^{n-1} (q^{1/2} + q^{-1/2})^{n-1}$. La norma $\|V_{C_n}(e^{2\pi i/5})\|$ è il rapporto aureo ϕ^{n-1} per il ciclo fondamentale sul cerchio a 5 elementi — connettendo la vaghezza con la geometria del pentagono regolare.

XIX.4 Invarianti di Vassiliev e la teoria delle derivazioni semantiche

Gli invarianti di Vassiliev (Vassiliev, 1990; Bar-Natan, 1995) sono una famiglia di invarianti di nodi di tipo finito: un invariante di nodi I è di tipo $\leq n$ (ordine di Vassiliev $\leq n$) se si annulla su tutti i nodi singolari con più di n punti doppi. Gli invarianti di Vassiliev di ordine 0 sono le costanti; di ordine 1 ci sono solo le costanti; di ordine 2 ci sono la costante e l'invariante c_2 (il secondo coefficiente del polinomio di Conway); di ordine 3 ci sono la costante, c_2 , e il numero di linking del nodo con suo stesso, e così via.

La teoria degli invarianti di Vassiliev ha un'algebra soggiacente ricca: gli invarianti di Vassiliev di ordine $\leq n$ formano uno spazio vettoriale V_n con $\dim V_n = 1, 1, 2, 4, 8, 12, \dots$ per $n = 0, 1, 2, 3, 4, 5, \dots$ (sequenza OEIS A007293). Il grafo degli invarianti di Vassiliev è la teoria delle «chord diagrams» e delle «weight systems» (Bar-Natan, 1995), con una connessione profonda con la teoria delle algebre di Lie via il teorema di Kontsevich.

Per i nodi semantici γ_C associati ai cicli del grafo di tolleranza, gli invarianti di Vassiliev danno una gerarchia di raffinamenti della classe di isotopia del nodo. Il primo invariante non banale $c_2(\gamma_C)$ (il secondo coefficiente del polinomio di Conway, che coincide con il coefficiente quadratico nel polinomio di Jones) misura l'«entanglement» del ciclo semantico: la quantità di

incroci nel diagramma minimale del nodo. Per cicli semantici planari, $c_2 = 0$. Per cicli su grafi non-planari di tolleranza, $c_2 \geq 1$.

Definizione XIX.2 (Invariante di Vassiliev del predicato vago). Per uno spazio di tolleranza $(X, \sim)$ con cicli indipendenti $C_1, \dots, C_\mu$ (μ = numero ciclomatico), l'invariante di Vassiliev del predicato P è il vettore: $\text{Vas}(P) = (c_2(\gamma_{\{C_1\}}), c_2(\gamma_{\{C_2\}}), \dots, c_2(\gamma_{\{C_\mu\}})) \in \mathbb{Z}^\mu$, dove $c_2(\gamma_{\{C_k\}})$ è il secondo coefficiente del polinomio di Conway del nodo semantico $\gamma_{\{C_k\}}$. L'invariante di Vassiliev del predicato è un invariante dello spazio di tolleranza rispetto a isomorfismi che preservano l'embedding in $\mathbb{R}^3$. Due predicati con il medesimo sorites barcode (stessa struttura omologica H_0, H_1) ma diversi invarianti di Vassiliev sono topologicamente distinti come sistemi di nodi semantici.

XIX.5 La relazione di skein semantica e la vaghezza come invariante di Reidemeister

Le mosse di Reidemeister (Reidemeister, 1927) sono le trasformazioni locali dei diagrammi di nodi che preservano la classe di isotopia. Ci sono tre tipi di mosse: Reidemeister I (aggiunta/rimozione di una ricciola), Reidemeister II (aggiunta/rimozione di una coppia di incroci opposti), Reidemeister III (scivolamento di un filo al di sotto/sopra di un incrocio). Il teorema di Reidemeister afferma che due diagrammi rappresentano nodi isotopi se e solo se sono connessi da una sequenza finita di mosse di Reidemeister.

Nel contesto semantico, le mosse di Reidemeister hanno un'interpretazione naturale come trasformazioni degli spazi di tolleranza:

La mossa di Reidemeister I corrisponde all'aggiunta o alla rimozione di un ciclo di lunghezza 1 nel grafo di tolleranza — il caso in cui un elemento è tollerante con sé stesso a livello «superiore» (una relazione di tolleranza che forma un cappio). Poiché i cappi sono sempre presenti per la riflessività, questa mossa è banale per gli spazi di tolleranza standard.

La mossa di Reidemeister II corrisponde all'aggiunta o alla rimozione di un «percorso di andata e ritorno» nel grafo di tolleranza: se $x_i \sim x_j \sim x_k \sim x_i$ forma un ciclo di lunghezza 3 e $x_i \sim x_k$ è un arco diretto, la mossa di Reidemeister II rimuove la coppia $(x_i \sim x_j, x_j \sim x_k)$ sostituendola con il singolo arco $(x_i \sim x_k)$. In termini di vaghezza, questa mossa corrisponde alla chiusura transitiva locale: l'aggiunta dell'arco diretto $x_i \sim x_k$ che «cortocircuita» il percorso di due passi.

La mossa di Reidemeister III corrisponde alla mossa di «scambio di triangolo»: se $x_i \sim x_j, x_j \sim x_k, x_i \sim x_k$ (un triangolo nel grafo di tolleranza), la mossa di Reidemeister III trasforma il ciclo (x_i, x_j, x_k, x_i) nel ciclo equivalente con un ordine di attraversamento diverso. In termini semantici, questa mossa corrisponde alla «commutatività delle catene di indistinguibilità»: l'ordine in cui si attraversano i tre passi non cambia il ciclo topologico.

Teorema XIX.2 (Invarianza del sorites barcode sotto mosse di Reidemeister semantiche). Le tre mosse di Reidemeister semantiche (aggiunta di cappi, chiusura transitiva locale, scambio di triangolo) preservano il sorites barcode del predicato vago: se $(X, \sim)$ e $(X, \sim')$ sono connessi da una sequenza di mosse di Reidemeister semantiche, allora $B(X, \sim) = B(X, \sim')$ (uguaglianza di sorites barcode). In particolare: il numero di risoluzione ε^* è invariante rispetto alle mosse di Reidemeister II e III (chiusura transitiva locale e scambio di triangolo), ma non rispetto alla mossa I (aggiunta di cappi che cambia il numero di elementi del dominio). Questa invarianza è più forte di quella sotto isomorfismi di spazi di tolleranza (che sono biiezioni che preservano la relazione $\sim$), e identifica una classe di equivalenza più larga di predicati con la stessa struttura di vaghezza topologica.

XIX.6 Il polinomio HOMFLY-PT e la classificazione completa dei predicati vaghi planari

Il polinomio HOMFLY-PT $P(v, z)$ (Freyd, Yetter, Hoste, Lickorish, Millett, Ocneanu, 1985; Przytycki-Traczyk, 1987) è un polinomio di Laurent a due variabili v e z che classifica i nodi in modo più fine del polinomio di Jones. La relazione di skein che lo definisce è:

$$v^{-1} P(L_+) - v P(L_-) = z P(L_0)$$

dove L_+ , L_- , L_0 sono tre diagrammi identici tranne in un intorno locale dove si trovano rispettivamente un incrocio positivo, un incrocio negativo, e un incrocio rimosso (due fili non intrecciati). Il polinomio di Jones è la specializzazione $v = q$, $z = q^{\{1/2\}} - q^{\{-1/2\}}$; il polinomio di Alexander è la specializzazione $v = 1$, $z = t^{\{1/2\}} - t^{\{-1/2\}}$.

Per i nodi semantici planari (che provengono da grafi di tolleranza planari), tutti i nodi sono isotopi al nodo banale (la curva chiusa senza incroci) e il polinomio HOMFLY-PT è il polinomio triviale $P_{\text{triviale}}(v, z) = 1$. La classificazione non banale emerge per grafi non-planari. Il teorema fondamentale di classificazione per nodi semantici di predicati planari è:

Teorema XIX.3 (Classificazione HOMFLY-PT dei predicati planari). Sia $(X, \sim)$ uno spazio di tolleranza con grafo $G(X, \sim)$ planare. Allora tutti i nodi semantici dei cicli indipendenti di $G(X, \sim)$ hanno polinomio HOMFLY-PT triviale $P = 1$. Due spazi di tolleranza planari $(X, \sim)$ e $(X, \sim')$ con lo stesso sorites barcode e lo stesso polinomio di Tutte $T_G(x, y) = T_{G'}(x, y)$ sono topologicamente equivalenti come sistemi di nodi semantici: i loro sistemi di invarianti nodali (Jones, HOMFLY-PT, Vassiliev) coincidono. Per spazi di tolleranza non-planari, il polinomio HOMFLY-PT del ciclo fondamentale del sorite (il ciclo ottenuto dalla composizione dei generatori di H_1) è un raffinamento del sorites barcode: due predicati con lo stesso barcode ma polinomio HOMFLY-PT diverso hanno strutture di vaghezza topologicamente distinte in $\mathbb{R}^3$.

XIX.7 Il gruppo fondamentale del complemento del nodo semantico e la π_1 della vaghezza

Il gruppo fondamentale $\pi_1(\mathbb{R}^3 \setminus \gamma_C)$ del complemento del nodo semantico γ_C in $\mathbb{R}^3$ è il gruppo di Wirtinger del nodo (Wirtinger, 1905). Per un diagramma di nodo con k incroci, il gruppo di Wirtinger ha una presentazione con k generatori (uno per ogni arco del diagramma) e k relazioni (una per ogni incrocio). Per il cerchio standard (nodo banale), $\pi_1(\mathbb{R}^3 \setminus \circ) = \mathbb{Z}$ (il gruppo degli interi, generato dal meridiano del toro standard).

Per i nodi semantici dei cicli del grafo di tolleranza, il gruppo di Wirtinger ha la seguente interpretazione: i generatori del gruppo corrispondono agli archi del ciclo semantico, e le relazioni corrispondono alle «interazioni topologiche» tra archi sovrapposti nell'immersione. Per cicli semantici planari, $\pi_1(\mathbb{R}^3 \setminus \gamma_C) = \mathbb{Z}$ per ogni ciclo — il complemento di una curva planare chiusa in $\mathbb{R}^3$ ha sempre gruppo fondamentale $\mathbb{Z}$. Per cicli non-planari, il gruppo di Wirtinger è non abeliano e cattura la struttura degli incroci del nodo.

Il abelianizzato del gruppo di Wirtinger è sempre $\mathbb{Z}$ (per nodi connessi): $\pi_1^{\text{ab}}(\mathbb{R}^3 \setminus \gamma_C) = \mathbb{Z}$. Questo riflette il fatto che $H_1(\mathbb{R}^3 \setminus \gamma_C; \mathbb{Z}) = \mathbb{Z}$ per ogni nodo — la topologia dell'omologia del complemento del nodo è universale. La differenza tra nodi diversi emerge nei gruppi di omologia superiori e nell'algebra del gruppo fondamentale non-abeliano.

Un risultato teorico importante è la connessione tra il gruppo di Wirtinger del nodo semantico e la struttura del gruppo fondamentale del complesso di Sorite $|S(X, \sim)|$:

Proposizione XIX.2 (Gruppo fondamentale del complesso di Sorite e del nodo semantico). Per uno spazio di tolleranza $(X, \sim)$ con grafo $G(X, \sim)$ connesso e ciclo fondamentale C (nel caso del sorite circolare, Capitolo 3), il gruppo fondamentale del complesso di Sorite $|S(X, \sim)|$ è: $\pi_1(|S(X, \sim)|) = \pi_1(|G(X, \sim)|) = \langle \text{generatori degli archi} \mid \text{relazioni delle facce} \rangle$ (presentazione di Seifert-van Kampen). La relazione con il gruppo di Wirtinger del nodo semantico γ_C è data dalla suriezione naturale $\pi_1(|S(X, \sim)|) \twoheadrightarrow \text{Wirt}(\gamma_C)$, che manda ogni generatore-arco del complesso di Sorite nel corrispondente generatore-arco di Wirtinger. Il nucleo della suriezione è il sottogruppo normale generato dalle relazioni aggiuntive delle facce 2-dimensionali del complesso di Sorite (i 2-simplessi) — relazioni che non esistono nel gruppo di Wirtinger del nodo 1-dimensionale.

SINTESI DELLE PARTI XVI–XIX: QUATTRO NUOVE PROSPETTIVE SULLA STRUTTURA FORMALE DELLA VAGHEZZA

Le quattro sezioni sviluppate in questo secondo supplemento approfondiscono la teoria degli spazi di tolleranza in direzioni radicalmente diverse, ciascuna delle quali rivela aspetti della struttura formale della vaghezza che le teorie precedenti non catturavano. La Parte XVI ha mostrato che il calcolo del numero di risoluzione ε^* è un'istanza di algebra min-plus (chiusura di Kleene tropicale), e che la varietà tropicale di tolleranza codifica la struttura degli assegnamenti coerenti in forma geometrica tropicale. Il deficit di transitività $\Delta_{\text{transit}}(\sim) = n(n-1)/2 - |E|$ misura la distanza della relazione di tolleranza dalla transitività completa — ed è proporzionale al numero di elementi aggiuntivi che l'algoritmo di Buchberger deve aggiungere alla base di Gröbner per completare l'ideale di tolleranza.

La Parte XVII ha dimostrato che il paradosso di Sorite nel senso dell'analisi non-standard ha costo infinitesimale ma genuinamente non-nullo: $\varepsilon^*(\sim_\eta) \sim \eta$ nell'estensione ${}^*\mathbb{R}$. Il Teorema di Compattezza per la vaghezza (XVII.2) stabilisce che il paradosso di Sorite emerge già su domini finiti e non richiede il continuo: la vaghezza è finitariamente compatta. La gerarchia di Hrbáček-Nelson dei predicati vaghi classifica i predicati in base all'ordine di infinitesimità del loro numero di risoluzione — una classificazione che va al di là della semplice distinzione vago/nitido.

La Parte XVIII ha collegato la vaghezza alla teoria ergodica: l'entropia di Kolmogorov-Sinai del sistema dinamico semantico è $\log(1/\varepsilon^*) + C$ — predicati più vaghi sono più caotici. Il Teorema di Birkhoff semantico dice che il predicato di lungo periodo di un parlante è il baricentro del poliedro $\mathcal{F}_{\{\varepsilon^*\}}$ — l'assegnamento ottimale f^* . Le misure SRB del sistema semantico si concentrano attorno a f^* con varianza $\varepsilon^{*2}/\lambda_{\text{eff}}$, confermando che gli elementi nella zona di confine sono i più variabili nelle classificazioni dei parlanti.

La Parte XIX ha introdotto la teoria dei nodi nel contesto degli spazi di tolleranza: i cicli del grafo di tolleranza, immersi in $\mathbb{R}^3$, producono nodi semantici il cui polinomio di Jones e HOMFLY-PT sono invarianti topologici del predicato più fini del sorites barcode. Per predicati planari, tutti i nodi semantici sono banali; per predicati non-planari emergono nodi non-triviali. La relazione di Reidemeister semantica fornisce una nuova classe di equivalenza tra predicati: due predicati connessi da mosse di Reidemeister hanno lo stesso barcode, ma la direzione inversa non vale — il polinomio HOMFLY-PT distingue predicati con lo stesso barcode ma struttura nodal-semantica diversa.

RIFERIMENTI AGGIUNTIVI PER LE PARTI XVI–XIX

- Mikhalkin G., «Enumerative tropical algebraic geometry in $\mathbb{R}^3$ », *Journal of the American Mathematical Society* 18(2), 313–377 (2005).
- Speyer D. e Sturmfels B., «The tropical Grassmannian», *Advances in Geometry* 4(3), 389–411 (2004).
- Sturmfels B., *Solving Systems of Polynomial Equations*, CBMS Regional Conference Series in Mathematics 97, American Mathematical Society (2002).
- Itenberg I., Mikhalkin G. e Shustin E., *Tropical Algebraic Geometry*, 2^a ed., Oberwolfach Seminars 35, Birkhäuser (2009).
- Maclagan D. e Sturmfels B., *Introduction to Tropical Geometry*, Graduate Studies in Mathematics 161, American Mathematical Society (2015).
- Jensen A., Markwig H. e Markwig T., «An algorithm for lifting points in a tropical variety», *Collectanea Mathematica* 59(2), 129–165 (2008).
- Lazarsfeld R. e Mustață M., «Convex bodies associated to linear series», *Annales Scientifiques de l'École Normale Supérieure* 42(5), 783–835 (2009).
- Kaveh K. e Khovanskii A. G., «Newton-Okounkov bodies, semigroups of integral points, graded algebras and intersection theory», *Annals of Mathematics* 176(2), 925–978 (2012).
- Tutte W. T., «A contribution to the theory of chromatic polynomials», *Canadian Journal of Mathematics* 6, 80–91 (1954).
- Buchberger B., *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*, Dissertazione, Universität Innsbruck (1965).
- Robinson A., *Non-Standard Analysis*, North-Holland (1966).
- Robinson A., «Non-standard analysis», *Proceedings of the Royal Academy of Sciences of Amsterdam A* 64, 432–440 (1961).
- Łoś J., «Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres», in *Mathematical Interpretation of Formal Systems*, North-Holland (1955).
- Nelson E., «Internal set theory: A new approach to nonstandard analysis», *Bulletin of the American Mathematical Society* 83(6), 1165–1198 (1977).
- Hrbáček K., «Axiomatic foundations for nonstandard analysis», *Fundamenta Mathematicae* 98(1), 1–19 (1978).
- Goldblatt R., *Lectures on the Hyperreals: An Introduction to Nonstandard Analysis*, Graduate Texts in Mathematics 188, Springer (1998).
- Cutland N. J. (ed.), *Nonstandard Analysis and its Applications*, London Mathematical Society Student Texts 10, Cambridge University Press (1988).
- Keisler H. J., *Elementary Calculus: An Infinitesimal Approach*, 2^a ed., Prindle, Weber & Schmidt (1986). Edizione open-access disponibile presso l'autore.
- Birkhoff G. D., «Proof of the ergodic theorem», *Proceedings of the National Academy of Sciences USA* 17(12), 656–660 (1931).
- von Neumann J., «Proof of the quasi-ergodic hypothesis», *Proceedings of the National Academy of Sciences USA* 18(1), 70–82 (1932).
- Kolmogorov A. N., «A new metric invariant of transitive dynamical systems and automorphisms in Lebesgue spaces», *Doklady Akademii Nauk SSSR* 119, 861–864 (1958).
- Sinai Ya. G., «On the notion of entropy of a dynamical system», *Doklady Akademii Nauk SSSR* 124, 768–771 (1959).
- Bowen R., *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Mathematics 470, Springer (1975).
- Ruelle D., «A measure associated with Axiom-A attractors», *American Journal of Mathematics* 98(3), 619–654 (1976).
- Poincaré H., *Les Méthodes Nouvelles de la Mécanique Céleste*, vol. 3, Gauthier-Villars (1899).
- Kac M., «On the notion of recurrence in discrete stochastic processes», *Bulletin of the American Mathematical Society* 53(10), 1002–1010 (1947).
- Walters P., *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics 79, Springer (1982).
- Cornfeld I. P., Fomin S. V. e Sinai Ya. G., *Ergodic Theory*, Grundlehren der mathematischen Wissenschaften 245, Springer (1982).

- Einsiedler M. e Ward T., *Ergodic Theory: With a View Towards Number Theory*, Graduate Texts in Mathematics 259, Springer (2011).
- Reidemeister K., «Elementare Begründung der Knotentheorie», *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* 5(1), 24–32 (1927).
- Alexander J. W., «Topological invariants of knots and links», *Transactions of the American Mathematical Society* 30(2), 275–306 (1928).
- Jones V. F. R., «A polynomial invariant for knots via von Neumann algebras», *Bulletin of the American Mathematical Society* 12(1), 103–111 (1985).
- Kauffman L. H., «State models and the Jones polynomial», *Topology* 26(3), 395–407 (1987).
- Freyd P., Yetter D., Hoste J., Lickorish W. B. R., Millett K. e Ocneanu A., «A new polynomial invariant of knots and links», *Bulletin of the American Mathematical Society* 12(2), 239–246 (1985).
- Przytycki J. H. e Traczyk P., «Invariants of links of Conway type», *Kobe Journal of Mathematics* 4(2), 115–139 (1987).
- Vassiliev V. A., «Cohomology of knot spaces», in *Theory of Singularities and its Applications*, *Advances in Soviet Mathematics* 1, 23–69, American Mathematical Society (1990).
- Bar-Natan D., «On the Vassiliev knot invariants», *Topology* 34(2), 423–472 (1995).
- Lickorish W. B. R., *An Introduction to Knot Theory*, Graduate Texts in Mathematics 175, Springer (1997).
- Adams C. C., *The Knot Book: An Elementary Introduction to the Mathematical Theory of Knots*, American Mathematical Society (2004).
- Cromwell P. R., *Knots and Links*, Cambridge University Press (2004).
- Murasugi K., *Knot Theory and its Applications*, Birkhäuser (1996).
- Kauffman L. H., *Knots and Physics*, 4^a ed., *Series on Knots and Everything* 1, World Scientific (2013).
- Birman J. S. e Lin X.-S., «Knot polynomials and Vassiliev's invariants», *Inventiones Mathematicae* 111(2), 225–270 (1993).
- Kontsevich M., «Vassiliev's knot invariants», *Advances in Soviet Mathematics* 16(2), 137–150 (1993).
- Appel K. e Haken W., «Every planar map is four colorable. Part I: Discharging», *Illinois Journal of Mathematics* 21(3), 429–490 (1977).
- Seifert H., «Über das Geschlecht von Knoten», *Mathematische Annalen* 110(1), 571–592 (1934).
- Whitney H., «2-isomorphic graphs», *American Journal of Mathematics* 55(1–4), 245–254 (1933).
- Kuratowski K., «Sur le problème des courbes gauches en topologie», *Fundamenta Mathematicae* 15(1), 271–283 (1930).

Capitolo finale sulla vaghezza linguistica: Struttura, Misura e Conseguenze

Il percorso compiuto nei capitoli precedenti ha costruito, passo dopo passo, un apparato matematico di notevole ampiezza e profondità per l'analisi del paradosso di Sorite. Spazi di tolleranza, grafi di tolleranza, complessi simpliciali, omologia persistente, teoria spettrale, barc-code di Sorite: ciascuno di questi strumenti è stato introdotto con motivazioni precise e dimostrato con rigore. È ora il momento di tornare al punto di partenza, quello che la premessa del saggio aveva anticipato come la vera posta in gioco: la vaghezza linguistica. Non come caso di studio esemplificativo del paradosso formale, ma come oggetto di indagine autonomo, la cui natura intrinseca — semantica, cognitiva, computazionale — il quadro matematico sviluppato permette finalmente di articolare con una precisione che le trattazioni filosofiche tradizionali non potevano raggiungere.

La vaghezza linguistica non è un difetto del linguaggio naturale, né una conseguenza della sua ambiguità o povertà espressiva. È una proprietà strutturale dei predicati: emerge dalla relazione tra il modo in cui i termini linguistici si applicano a oggetti del mondo e la struttura dell'indistinguibilità percettiva e concettuale degli oggetti stessi. Questa sezione dimostra che la vaghezza linguistica è esattamente la topologia non banale del complesso di Sorite associato al predicato, e che la sua misura quantitativa è il numero di risoluzione ε^* definito nel Capitolo 5. L'intera apparecchiatura matematica sviluppata nel saggio si rivela, in questo capitolo finale, come una teoria rigorosa della vaghezza linguistica, non del semplice paradosso logico che ne è la manifestazione più nota.

1 Che cosa è la vaghezza linguistica: una definizione strutturale

La vaghezza linguistica di un predicato P è, nella terminologia del presente saggio, l'insieme delle proprietà topologiche non banali del complesso di Sorite $S(X, \sim_P)$ associato a P sul suo dominio di applicazione X , dove $\sim_P$ è la relazione di tolleranza indotta da P su X attraverso la struttura di indistinguibilità percettiva e concettuale degli elementi di X rispetto a P .

Questa definizione strutturale — la vaghezza come topologia — ha il grande vantaggio di essere operativa: essa si può calcolare, misurare, confrontare tra predicati diversi, e studiare nelle sue variazioni al variare della granularità del dominio. Non è,

invece, una definizione descrittiva nel senso delle teorie del significato classiche, che cercano di catturare intuitivamente cosa voglia dire per un termine essere vago. La definizione strutturale è compatibile con tutte le teorie filosofiche della vaghezza discusse nei Capitoli 1.3 e 4.5 — epistemicismo, supervalutazione, logiche fuzzy — nel senso che ciascuna di quelle teorie può essere riformulata come una scelta di come rispondere alla topologia non banale del complesso di Sorite, non come una negazione della sua esistenza.

Il complemento della vaghezza, nella stessa terminologia, è la nitidezza: un predicato P è nitido se e solo se il complesso $S(X, \sim_P)$ è disconnesso nel senso che le istanze paradigmaticamente vere e quelle paradigmaticamente false di P appartengono a componenti connesse distinte di $G(X, \sim_P)$. Il Corollario 1.1 del Capitolo 4 caratterizza questo come la condizione necessaria e sufficiente per l'esistenza di un assegnamento coerente con contrasto. La nitidezza non richiede che il confine tra le componenti sia localizzato in un singolo punto — possono esistere elementi in zone intermedie — ma soltanto che le componenti contenenti il minimo e il massimo del dominio ordinato siano disgiunte nel grafo di tolleranza.

Un predicato P su un dominio $(X, \leq, d)$ — dove $\leq$ è l'ordinamento totale e d è la distanza che genera la tolleranza $\sim_P$ — è *vago* in senso strutturale se e solo se $G(X, \sim_P)$ è connesso. È *nitido* se e solo se $G(X, \sim_P)$ è disconnesso con $\min(X)$ e $\max(X)$ in componenti distinte. Il grado di vaghezza è quantificato dal numero di risoluzione $\varepsilon^*(\sim_P) = 1/L_{\text{mono_min}}$, dove $L_{\text{mono_min}}$ è la lunghezza del cammino monotono minimo nel DAG $G^{\rightarrow}(X, \sim_P)$. Quando $\varepsilon^*(\sim_P) \rightarrow 0$, il predicato diventa quasi-nitido; quando $\varepsilon^*(\sim_P) = 1$, la transizione è concentrata in un unico passo nel cammino monotono, il che corrisponde alla vaghezza massima compatibile con la struttura metrica del dominio.

Il riferimento alla distanza d merita una precisazione. In linea di principio, la vaghezza dipende dalla struttura di indistinguibilità $\sim_P$, non dalla metrica d . La metrica è solo uno strumento per generare $\sim_P$: si pone $x \sim_P y$ se $d(x, y) \leq r$ per una soglia $r > 0$ (la JND, nel caso della psicofisica). Predicati diversi sullo stesso dominio metrico possono avere soglie r diverse e quindi strutture di tolleranza diverse. Anche la scelta della metrica d può essere diversa per predicati diversi sullo stesso dominio: la distanza rilevante per il predicato "alto" sulla statura non è la stessa distanza rilevante per il predicato "giovane" sull'età, perché le strutture percettive e contestuali di questi due predicati sono diverse.

2 Predicati gradabili e struttura metrica dell'indistinguibilità

La classe di predicati vaghi più studiata in linguistica formale è quella dei predicati gradabili (gradable adjectives, nella terminologia inglese): aggettivi come "alto", "caldo", "veloce", "ricco", "vecchio", "grande" che ammettono comparativi («più alto di»), superlativi («il più alto»), costruzioni con avverbi di grado («molto alto», «abbastanza alto», «piuttosto alto»). La teoria semantica standard dei predicati gradabili — quella di Cresswell (1977), Kennedy (1999), Kennedy e McNally (2005) — li tratta come funzioni da individui a gradi su una scala ordinata: «alto» denota una funzione che mappa ogni individuo al suo grado di altezza, e il predicato vero/falso emerge dalla comparazione di questo grado con uno standard contestuale.

Nel quadro del presente saggio, la scala ordinata su cui i gradi sono misurati corrisponde al dominio $(X, \leq)$. La struttura di tolleranza $\sim_P$ corrisponde alla relazione di indistinguibilità percettiva lungo la scala: due individui x e y con gradi $d(x)$ e $d(y)$ tali che $|d(x) - d(y)| \leq r$ (dove r è la JND per il predicato P) sono nella relazione $x \sim_P y$. Il numero di risoluzione $\varepsilon^*(\sim_P)$ misura quanto strettamente il predicato può discriminare tra il polo positivo e quello negativo della scala, tenuto conto della struttura di indistinguibilità.

Questa corrispondenza non è solo nominale: permette di derivare predizioni quantitative verificabili. Per il predicato "alto" applicato alla statura maschile adulta in Italia, si possono stimare i parametri rilevanti. Il range della statura adulta maschile è approssimativamente 150–210 cm, con una distribuzione normale centrata intorno a 176 cm con deviazione standard di circa 7 cm. La JND visiva per la statura quando le persone sono affiancate è stimata nell'ordine di 5–10 mm, come indicato nella Sezione 2.1. Si pone $r = 7 \text{ mm} = 0,7 \text{ cm}$ come stima della JND. Il dominio discretizzato ha $n = (210 - 150) / 0,7 \approx 86$ elementi. La tolleranza è di larghezza uniforme 1 rispetto alla discretizzazione, il che significa $bw_{\leq} = 1$ e $L_{\text{mono_min}} = n - 1 = 85$. Il numero di risoluzione è:

$$\varepsilon^*(\sim_{\text{alto}}) = 1 / L_{\text{mono_min}} = 1 / 85 \approx 0,0118$$

Questo valore significa che qualunque assegnamento coerente del valore di verità «è alto» lungo la scala della statura, compatibile con la JND di 0,7 cm, deve necessariamente variare di almeno $1/85 \approx 1,18\%$ del range $[0, 1]$ tra ciascun passo lungo il cammino monotono. Se si usa la funzione lineare ottimale (che realizza il minimo ε^*), ogni millimetro di altezza aggiuntivo corrisponde a un aumento del valore di verità di circa $1/85 \approx 1,18\%$.

Un'analogia illuminante emerge confrontando predicati sulla stessa scala con JND diverse. Per il predicato "molto alto" (diciamo, con standard contestuale più elevato), la JND percettiva potrebbe essere più grande — diciamo 1 cm — perché si valutano differenze rispetto a un prototipo più estremo. In questo caso $n = (210 - 150) / 1 = 60$ e $\varepsilon^*(\sim_{\{\text{molto alto}\}}) = 1/59 \approx 1,69\%$. Il predicato "molto alto" è più vago di "alto" in senso strutturale: ha un numero di risoluzione più grande, e la transizione dal polo negativo al polo positivo è distribuita su meno passi. Questo è controintuitivo rispetto all'uso ordinario, ma riflette correttamente la struttura percettiva: differenze di altezza al di sopra di una certa soglia sono meno finemente discriminabili, non più.

3 La geometria della zona di confine

Uno degli aspetti della vaghezza linguistica più discussi nella letteratura filosofica è la zona di confine (borderline zone, penumbra): la regione del dominio in cui il predicato non si applica chiaramente né come vero né come falso. Nella teoria degli spazi di tolleranza, la zona di confine ha una caratterizzazione topologica precisa che ne rende misurabile sia l'ampiezza sia la struttura.

Per uno spazio di tolleranza $(X, \sim_P, \leq)$ con $G(X, \sim_P)$ connesso, la zona di confine del predicato P rispetto a un assegnamento ottimale $f: X \rightarrow [0, 1]$ è l'insieme:

$$B_{\delta}(f) = \{ x \in X : \delta \leq f(x) \leq 1 - \delta \}$$

per un parametro $\delta \in (0, 1/2]$ che determina il «grado di chiarezza» richiesto per l'appartenenza alla zona di confine. Gli elementi con $f(x) < \delta$ sono «chiaramente falsi»; quelli con $f(x) > 1 - \delta$ sono «chiaramente veri». La zona di confine $B_{\delta}(f)$ contiene esattamente gli elementi per cui il valore di verità è «significativamente lontano dai poli».

Per la funzione gradino ottimale costruita nella Parte 1 della dimostrazione del Teorema 3, la zona di confine $B_{\{1/L\}}(f)$ — con soglia pari a $\varepsilon^* = 1/L_{\text{mono_min}}$ — è l'insieme degli elementi $x \in X$ che si trovano tra il primo e l'ultimo nodo del cammino monotono minimo in senso stretto, escludendo i nodi con valori di verità esattamente 0 o 1. La dimensione di questa zona dipende dalla struttura del grafo e non è in generale determinata soltanto da $L_{\text{mono_min}}$.

Esiste un legame preciso tra l'ampiezza della zona di confine e il gap spettrale μ_1 del laplaciano normalizzato del grafo $G(X, \sim_P)$, formalizzato dalla disuguaglianza di Cheeger (Teorema 8 del Capitolo 8). La disuguaglianza afferma che:

$$h^2 / 2 \leq \mu_1 \leq 2h$$

dove h è la costante di Cheeger isoperimetrica e μ_1 è il primo autovalore non nullo del laplaciano. Un gap spettrale μ_1 piccolo (h piccolo) indica la presenza di un «collo di bottiglia» nel grafo: un sottoinsieme S di vertici con volume grande ma bordo piccolo. Semanticamente, questo collo di bottiglia corrisponde alla zona di confine del predicato: la regione del dominio in cui pochi archi di tolleranza «attraversano» il confine tra la zona chiaramente vera e quella chiaramente falsa. Predicati con zona di confine pronunciata (psicologicamente «difficili» da classificare) hanno μ_1 piccolo; predicati con transizione netta (psicologicamente «facili») hanno μ_1 grande.

Questa connessione tra il gap spettrale e la difficoltà psicologica di classificazione è verificabile empiricamente. Esperimenti di tempo di reazione su classificazioni di stimoli ambigui — studi classici di Barsalou (1987) e di Hampton (1979) sulle categorie prototipiche — mostrano che gli stimoli nella zona di confine richiedono tempi di reazione significativamente più lunghi. Nella formalizzazione topologica, questo corrisponde al fatto che la passeggiata casuale su $G(X, \sim_P)$ raggiunge l'equilibrio più lentamente (il tempo di mixing è inversamente proporzionale a μ_1) negli elementi della zona di confine.

4 La dipendenza contestuale della vaghezza

Un fenomeno fondamentale della vaghezza linguistica, ampiamente documentato in linguistica e psicologia cognitiva, è la dipendenza contestuale dei predicati vaghi: lo stesso predicato applica standard diversi in contesti diversi, e questi standard influenzano la collocazione della zona di confine e, di conseguenza, il valore di verità degli enunciati nella zona di confine stessa. La domanda «Gianni è alto?» ha risposte diverse a seconda che si stia discutendo di una squadra di basket, di una commissione accademica, o di una piscina per bambini. Kennedy e McNally (2005), Barker (2002), e Rothschild e Segal (2009) hanno sviluppato teorie semantiche formali della contestualità dei predicati gradabili.

Nel quadro degli spazi di tolleranza, la dipendenza contestuale si traduce in una famiglia di spazi di tolleranza $(X, \sim_P^c)$ indicizzata dal contesto c , dove la relazione $\sim_P^c$ varia al variare del contesto. In senso formale, il contesto determina:

(a) la soglia di JND r^c , che dipende dal contesto perché l'attenzione dei parlanti è focalizzata diversamente — in un contesto in cui l'altezza è una proprietà saliente (squadra di basket), la JND è più piccola (i parlanti sono più attenti alle differenze di altezza);

(b) il sottodominio $X^c \subseteq X$ di confronto: la classe di riferimento rispetto alla quale il predicato è valutato — gli atleti di basket, i professori universitari, i bambini in età prescolare;

(c) lo standard contestuale θ^c : il «grado soglia» al di sopra del quale il predicato è vero, nella semantica dei gradi.

Il quadro unificante della teoria degli spazi di tolleranza permette di trattare (a) e (b) direttamente — la relazione $\sim_P^c$ è determinata da r^c e X^c — ma non (c) direttamente. Lo standard contestuale θ^c è codificato nella scelta delle condizioni agli estremi dell'assegnamento: invece di richiedere $f(\min(X^c)) = 0$ e $f(\max(X^c)) = 1$, si richiedono $f(\min(X^c)) = 0$ e $f(\theta^c) = 1$ (o $f(\theta^c) = 1/2$, secondo diverse convenzioni). Questo non cambia il calcolo di ε^* , ma modifica la funzione di membership ottimale, spostandone il «punto di flesso» verso θ^c .

Un risultato importante sulla contestualità è la stabilità del numero di risoluzione rispetto a perturbazioni del contesto, che segue direttamente dal Teorema 12 del Capitolo 9 sulla stabilità dei diagrammi di persistenza. Se due contesti c e c' differiscono per una perturbazione della metrica di ordine δ (nel senso che le distanze tra tutti gli elementi cambiano di al più δ), allora i sorites barcode dei due spazi di tolleranza differiscono per una distanza bottleneck di al più δ :

$$d_B(B(\sim_P^c), B(\sim_P^{c'})) \leq \delta$$

In termini linguistici, una piccola variazione contestuale produce una piccola variazione del sorites barcode e, di conseguenza, del numero di risoluzione ε^* . I predicati vaghi sono strutturalmente stabili rispetto a piccole variazioni contestuali — il che spiega perché la comunicazione linguistica con predicati vaghi sia possibile nonostante la variabilità contestuale: i parlanti condividono approssimativamente la stessa struttura topologica del predicato, anche se i parametri precisi variano.

5 Vaghezza di ordine superiore: la vaghezza della zona di confine

Uno dei problemi più profondi nella letteratura sulla vaghezza è la vaghezza di ordine superiore (higher-order vagueness o super-vagueness): non solo il confine tra il vero

e il falso è vago, ma il confine tra il chiaramente vero e il vagamente vero (la penumbra della penumbra) è anch'esso vago, e così via all'infinito. Questo fenomeno produce un regresso che mina alcune teorie della vaghezza — in particolare quelle che individuano una zona di confine definita e netta — e costituisce una delle sfide più gravi alla formalizzazione.

Nel quadro matematico del presente saggio, la vaghezza di ordine superiore corrisponde alla struttura gerarchica della filtrazione di Sorite. La filtrazione è la famiglia $\{S(X, \sim_r)\}_{r \geq 0}$ indicizzata dal parametro di scala r . Il sorites barcode — il diagramma di persistenza PD_0 della filtrazione — ha una struttura multiscala: le barre con lunga persistenza corrispondono alle componenti topologicamente stabili (che sopravvivono a variazioni di r), mentre le barre brevi corrispondono a strutture fragili. La vaghezza di primo ordine è codificata nella persistenza della barra principale di PD_0 (la componente che connette $\min(X)$ a $\max(X)$). La vaghezza di ordine superiore è codificata nella struttura delle barre secondarie: barre con persistenza intermedia corrispondono a «confini dentro il confine», cioè a suddivisioni della zona di confine in sottozone distinguibili su scale diverse.

Formalmente, sia (X, d) lo spazio metrico del dominio di P e sia r_0 il valore di r al quale la componente principale di $S(X, \sim_r)$ nasce (cioè al quale $\min(X)$ e $\max(X)$ si trovano per la prima volta nella stessa componente connessa, detto anche il valore di «coalescenza» nella filtrazione). La barra principale di PD_0 ha coordinate di nascita 0 ($\min(X)$ e $\max(X)$ sono nell'insieme fin dall'inizio nella filtrazione) e morte r_0 (il momento in cui perdono la loro separazione topologica). Le barre di PD_0 con coordinate di nascita $r_1 \in (0, r_0)$ corrispondono a componenti intermedie: a questi raggi di tolleranza, alcune parti del dominio si «coalescono» tra loro prima della coalescenza globale. Queste componenti intermedie corrispondono a sottoregioni del dominio che sono topologicamente «simili» a scala r_1 — cioè a «prototipi locali» del predicato, regioni in cui gli elementi sono indistinguibili tra loro ma distinguibili dagli elementi di altre regioni.

La struttura della vaghezza di ordine superiore è quindi la struttura multiscala del diagramma di persistenza PD_0 : la gerarchia di aggregazioni topologiche all'interno del dominio del predicato, ciascuna delle quali corrisponde a un livello di vaghezza diverso. Un predicato con molte barre di lunga persistenza in PD_0 ha molti livelli di vaghezza sovrapposti — la sua struttura di confine è «profonda». Un predicato con una sola barra

dominante e molte barre trascurabili ha una struttura di confine «piatta»: esiste sostanzialmente un solo livello di vaghezza, e la vaghezza di ordine superiore è trascurabile rispetto alla vaghezza di primo ordine.

La vaghezza di ordine superiore può essere quantificata attraverso la complessità del diagramma di persistenza, misurata, ad esempio, dall'entropia della distribuzione delle persistenze. Sia $\{(b_i, d_i)\}_i$ l'insieme dei punti del diagramma di persistenza PD_0 con $b_i < d_i$ (barre finite). La persistenza di ciascuna barra è $p_i = d_i - b_i$. L'entropia normalizzata delle persistenze è:

$$H(PD_0) = - \sum_i (p_i / \sum_j p_j) \cdot \log(p_i / \sum_j p_j)$$

Un predicato con vaghezza di ordine superiore pronunciata ha $H(PD_0)$ grande (molte barre con persistenza comparabile); un predicato con vaghezza essenzialmente di primo ordine ha $H(PD_0)$ piccola (una barra dominante e molte barre trascurabili). L'entropia del diagramma di persistenza è un nuovo invariante della vaghezza linguistica che non ha precedenti, per quanto è noto, nella letteratura filosofica o linguistica sulla vaghezza.

6 Vaghezza e semantica delle scale: scale aperte, chiuse e parzialmente chiuse

La linguistica formale distingue tra diversi tipi di scale associate ai predicati gradabili, a seconda della struttura dei punti estremi. Kennedy e McNally (2005) identificano quattro tipi: scale totalmente aperte (come «alto», «caldo»), scale totalmente chiuse (come «pieno», «vuoto»), scale aperte superiormente (come «macchiato»), e scale chiuse superiormente (come «asciutto»). Questa tassonomia ha conseguenze importanti per il comportamento dei predicati con avverbi di grado («completamente», «quasi», «per metà»).

Nel quadro degli spazi di tolleranza, la distinzione tra scale aperte e chiuse corrisponde alla struttura del dominio X vicino agli estremi. Per una scala chiusa inferiormente (il caso di «pieno» e «vuoto»: c'è un grado zero assoluto di riempimento), il valore minimo di X ha una JND molto piccola — la discriminazione vicino allo zero assoluto è fine — mentre il valore massimo ha una JND più grande. Per una scala aperta (il caso di «alto»), la JND è approssimativamente costante (uniforma la struttura di tolleranza) o cresce verso gli estremi (la legge di Weber-Fechner).

Formalmente, per una scala aperta con JND proporzionale all'intensità del segnale (legge di Weber-Fechner), la relazione di tolleranza $\sim_P^r$ è:

$$x \sim_P^r y \iff |x - y| / \max(x, y) \leq r$$

per un parametro $r > 0$. Questa tolleranza non è di larghezza uniforme: la larghezza della banda di tolleranza varia proporzionalmente all'intensità. Il numero di risoluzione ε^* per questa struttura dipende dalla forma specifica del dominio e dalla soglia r , e non si riduce alla formula semplice $\varepsilon^* = k/(n - 1)$ delle tolleranze uniformi.

Si calcola ε^* per una scala logaritmica, che è la struttura naturale per la percezione delle intensità fisiche. Sia $X = \{x_1 < x_2 < \dots < x_n\}$ con $x_i = e^{\{(i-1) \cdot h\}}$ per un passo $h > 0$ (scala geometrica). La tolleranza di Weber è $x_i \sim_r x_j$ se $|x_i - x_j| / x_j \leq r$ (con $i \leq j$), il che equivale a $e^{\{[i-j] \cdot h\}} \leq 1 + r$, ossia $|i-j| \leq \log(1+r)/h$. Denotando $k = \lfloor \log(1+r)/h \rfloor$, la tolleranza di Weber su scala logaritmica è equivalente a una tolleranza di larghezza uniforme k rispetto all'indice i . Il numero di risoluzione è quindi $\varepsilon^*(\sim_P^r) = k/(n-1) = \lfloor \log(1+r)/h \rfloor / (n-1)$, la stessa formula delle tolleranze uniformi ma con k determinato dal parametro psicofisico r e dal passo di discretizzazione h .

Questo risultato ha una implicazione importante: per predicati con legge di Weber-Fechner, la vaghezza (misurata da ε^*) è determinata esclusivamente dal rapporto $\log(1+r)/h$, che è indipendente dall'intensità assoluta. Due predicati con la stessa costante di Weber r ma su scale con diverso passo di discretizzazione hanno numeri di risoluzione diversi. Questo spiega perché la vaghezza del predicato «è alto» applicato a stature umane sia diversa dalla vaghezza del predicato «è alto» applicato ad edifici, anche se la soglia di indistinguibilità relativa (la costante di Weber) è la stessa: il passo di discretizzazione rilevante per gli edifici è diverso da quello per le stature umane.

7 Vaghezza multidimensionale: il caso dei predicati valutativi

I predicati semanticamente più interessanti e cognitivamente più complessi non sono quelli unidimensionali ("alto", "caldo") ma quelli genuinamente multidimensionali: predicati valutativi come "bello", "intelligente", "buono", "giusto", "sano", "in forma". Questi predicati dipendono simultaneamente da molteplici dimensioni di variazione, ciascuna delle quali è dotata di una propria struttura di indistinguibilità, e le dimensioni non si ordinano in modo univoco su una scala unica.

Il Capitolo 7 ha sviluppato la teoria degli spazi di tolleranza prodotto per predicati multidimensionali e calcolato i gruppi di omologia delle griglie bidimensionali. Il risultato fondamentale — che H_1 della griglia $m \times n$ ha rango $(m-1)(n-1)$ — ha conseguenze profonde per la vaghezza dei predicati valutativi. Ogni ciclo indipendente in H_1 corrisponde a un «paradosso ciclico locale» nel dominio bidimensionale del predicato: quattro istanze del predicato che formano un ciclo di indistinguibilità nel quale non è possibile assegnare in modo coerente i valori «vero» e «falso». I $(m-1)(n-1)$ cicli indipendenti sono $(m-1)(n-1)$ paradossi di Sorite sovrapposti e indipendenti.

Per il predicato "è intelligente", la ricerca psicologica ha identificato tra 5 e 9 dimensioni di intelligenza cognitive relativamente indipendenti (Gardner 1983, Sternberg 1985), anche se il dibattito sulla struttura fattoriale dell'intelligenza è ancora aperto. Si consideri, a titolo di esempio, un modello semplificato con soltanto 3 dimensioni: intelligenza logico-matematica I_L , intelligenza linguistica I_V (verbale), e intelligenza spaziale I_S . Ciascuna dimensione è quantificata su una scala da 1 a 10 con tolleranza di larghezza 1. Il dominio è la griglia 3-dimensionale $10 \times 10 \times 10$ con 1000 elementi.

Dalla Proposizione 16 del Capitolo 7, per la griglia cubica d-dimensionale di lato $n = 10$ con $d = 3$:

$$\text{rk}(H_1) = |E| - |V| + 1$$

con $|V| = 10^3 = 1000$ e $|E| = 3 \times (10-1) \times 10^2 = 2700$ (archi in ciascuna delle 3 direzioni). Dunque:

$$\text{rk}(H_1) = 2700 - 1000 + 1 = 1701$$

Ci sono 1701 ostruzioni omologiche indipendenti agli assegnamenti coerenti del predicato "è intelligente" su questo dominio tridimensionale. Ognuna di queste ostruzioni corrisponde a un quadrato elementare nella griglia, cioè a 4 individui che si trovano nelle quattro celle adiacenti di una faccia della griglia tridimensionale, e che non possono ricevere in modo coerente valori di verità distinti.

Il numero di risoluzione per questo predicato, con ordinamento lessicografico (I_L, I_V, I_S), è:

$$\varepsilon^*(\sim_{\{\text{intell}\}}) = 1 / (10 + 10 + 10 - 3) = 1/27 \approx 0,0370$$

Il valore $1/27$ è significativamente più piccolo di $1/9 \approx 0,111$, che sarebbe il numero di risoluzione di ciascuno dei tre predicati unidimensionali componenti (intelligenza logica

da sola, intelligenza verbale da sola, intelligenza spaziale da sola, ciascuna con 10 livelli). Questo riflette la struttura matematica descritta nella Proposizione 15.1: il predicato multidimensionale è meno vago (in senso graduato) del peggiore dei predicati unidimensionali componenti, perché il cammino monotono nella griglia tridimensionale può «saltare» attraverso le dimensioni, sfruttando la struttura prodotto.

Il risultato ha una conseguenza praticamente importante: i predicati valutativi multidimensionali, per quanto siano percepiti come «particolarmente soggettivi» o «difficili da classificare», non sono necessariamente più vaghi dei predicati unidimensionali nel senso del numero di risoluzione ε^* . Ciò che è maggiore, nei predicati multidimensionali, è il numero di ostruzioni omologiche H_1 — il numero di «paradossi locali» sovrapposti — non il grado minimo di incoerenza da accettare per ottenere una classificazione con contrasto.

8 Predicati di colore e struttura ciclica della vaghezza

I predicati di colore occupano una posizione speciale nella teoria della vaghezza, per tre ragioni convergenti. In primo luogo, la percezione del colore è uno degli ambiti in cui le strutture psicofisiche di indistinguibilità sono più accurate e meglio documentate sperimentalmente. In secondo luogo, la struttura dello spazio percettivo del colore ha topologia non banale: lo spazio CIE $L^*a^*b^*$ che modella la percezione cromatica umana è localmente euclidea ma globalmente non semplicemente connesso lungo la dimensione della tinta (hue). In terzo luogo, la categorizzazione cromatica nei sistemi linguistici del mondo è un banco di prova privilegiato per le teorie del relativismo linguistico e del determinismo dell'universo del discorso.

La struttura dell'hue circle, introdotta nell'Esempio 11 del Capitolo 6, è un cerchio S^1 parametrizzato dall'angolo di tinta $\theta \in [0^\circ, 360^\circ)$, con $\theta = 0^\circ$ che corrisponde al rosso puro e $\theta = 360^\circ$ che si identifica con $\theta = 0^\circ$. La JND cromatica lungo l'hue circle è approssimativamente $1-3^\circ$ nelle regioni di maggiore sensibilità (verde-giallo intorno a 555 nm, dove la JND è minima) e $5-10^\circ$ nelle regioni di minore sensibilità (rosso, viola).

Si discretizza l'hue circle con n passi angolari uguali di ampiezza $\Delta\theta = 360^\circ/n$, ottenendo n elementi $X_n = \{\theta_1, \theta_2, \dots, \theta_n\}$ con $\theta_k = (k-1) \cdot 360^\circ/n$. La tolleranza di adiacenza è $\theta_i \sim_1 \theta_j$ se $d_{\text{circ}}(\theta_i, \theta_j) \leq \Delta\theta$ (dove d_{circ} è la distanza circolare). Per $n = 360$ (un elemento per grado), si ha $\Delta\theta = 1^\circ$ e il grafo $G(X_{360}, \sim_1)$ è il ciclo C_{360} .

Il complesso di Sorite $S(X_{360}, \sim_1)$ è 1-dimensionale (nessun simpleso di dimensione ≥ 2 , poiché i vicini di ogni nodo sul ciclo non sono adiacenti tra loro), e ha:

$$H_0(S(X_{360}, \sim_1); \mathbb{Z}) \cong \mathbb{Z}, \quad H_1(S(X_{360}, \sim_1); \mathbb{Z}) \cong \mathbb{Z}$$

Il generatore di H_1 è il ciclo fondamentale dell'hue circle. L'ostruzione omologica in H_1 corrisponde all'impossibilità di assegnare a ogni tinta un valore di verità del predicato «è caldo come colore» in modo che:

(a) l'assegnamento sia coerente con la JND cromatica (colori indistinguibili ricevono lo stesso valore),

(b) l'assegnamento sia non costante (almeno alcune tinte calde e alcune tinte fredde ricevono valori diversi).

Il Teorema 4 del Capitolo 6 dimostra che queste due condizioni sono incompatibili: qualunque assegnamento coerente del predicato «è caldo come colore» sull'hue circle deve essere costante. Il paradosso è puramente strutturale e non dipende dalla scelta specifica delle tinte «calde» e «fredde».

Il numero di risoluzione ciclico per il ciclo C_{360} è, dalla formula del Teorema 5:

$$\varepsilon^*_{\text{cicl}} = 1 / \lfloor 360/2 \rfloor = 1/180 \approx 0,0056$$

Questo valore molto piccolo indica che l'hue circle è «quasi ciclicamente risolubile»: con un grado di incoerenza di soli 0,56% per passo, è possibile assegnare valori non costanti alle tinte cromatiche. In pratica, questo corrisponde al fatto che i parlanti usano predicati di colore in modo contestualmente coerente su large porzioni dell'hue circle, anche se la struttura ciclica globale è irrisolvibile.

Un risultato interessante riguarda l'effetto della larghezza della tolleranza sulla struttura omologica. Per la tolleranza di larghezza 2 sul ciclo C_n (ogni elemento è tollerante con i due vicini a destra e i due a sinistra), la Proposizione 14 del Capitolo 6 afferma che $H_1 = \mathbb{Z}$ se n non è divisibile per 3, e $H_1 = 0$ se n è divisibile per 3. Per $n = 360$ (divisibile per 3), la tolleranza di larghezza 2 elimina l'ostruzione omologica: i triangoli creati dalla larghezza maggiore «riempiono» il ciclo. Semanticamente, questo significa che una JND di 2° (invece di 1°) rende il predicato «è caldo come colore» sull'hue circle risolubile per assegnamento coerente non costante. La vaghezza ciclica dipende in modo non monotono dalla larghezza della JND.

9 Semantica composizionale e vaghezza

Un aspetto della vaghezza linguistica che non è stato trattato nei capitoli precedenti, ma che merita un'analisi nell'ottica della teoria sviluppata, è la semantica composizionale: il modo in cui la vaghezza dei predicati elementari si propaga alla vaghezza di espressioni complesse formate da operatori logici (negazione, congiunzione, disgiunzione) e quantificatori (tutti, alcuni, la maggior parte).

Nella semantica classica vero-condizionale (Montague, 1973), il significato di un'espressione complessa è una funzione del significato delle sue parti. Per espressioni contenenti predicati vaghi, questa compositività si scontra con la non-bivalenza: il valore di verità di «Gianni è alto e Bruno è giovane» non è determinato dal solo fatto che «Gianni è alto» è vero e «Bruno è giovane» è vagamente vero.

Nel quadro degli spazi di tolleranza, la semantica composizionale si tratta mediante i morfismi di spazi di tolleranza (Definizione 7 del Capitolo 3) e la funzorialità del complesso di Sorite (Proposizione 3). Se P e Q sono due predicati vaghi sui domini $(X_P, \sim_P)$ e $(X_Q, \sim_Q)$, il predicato congiunto « P e Q » opera sul dominio prodotto $(X_P \times X_Q, \sim_{\{P \times Q\}})$ con la tolleranza prodotto della Definizione 15 del Capitolo 7. Il complesso di Sorite del predicato congiunto è il complesso di Sorite del prodotto degli spazi di tolleranza.

Dalla teoria della griglia (Teorema 6 e Proposizione 16), il numero di ostruzioni omologiche del predicato congiunto è significativamente maggiore della somma delle ostruzioni dei predicati componenti. Per due predicati unidimensionali con m e n livelli rispettivamente, i predicati componenti hanno 0 cicli (nessuna ostruzione in H_1 , poiché le catene lineari sono contraibili), ma il predicato congiunto « P e Q » ha $(m-1)(n-1)$ cicli. La composizione logica crea nuova topologia non banale che non era presente nei componenti — questo è un fenomeno di emergenza topologica nella semantica composizionale.

Per la negazione, il morfismo di negazione $\varphi_{\{\text{neg}\}}: (X, \sim_P, \leq) \rightarrow (X, \sim_P, \geq)$ inverte l'ordinamento ma preserva la relazione di tolleranza. Il complesso di Sorite è isomorfo (i vertici e gli archi sono gli stessi), ma le condizioni agli estremi si invertono: il minimo dell'ordinamento originale diventa il massimo del predicato negato. Il numero di risoluzione è invariante: $\varepsilon^*(\sim_{\{\neg P\}}) = \varepsilon^*(\sim_P)$, perché $L_{\text{mono_min}}$ è calcolato nel DAG orientato, che inverte insieme all'ordinamento. Semanticamente, la negazione di un

predicato vago ha lo stesso grado di vaghezza del predicato originale — risultato che è intuitivo ma che qui emerge come conseguenza strutturale della teoria.

Per la disgiunzione, il morfismo di disgiunzione produce una struttura di tolleranza che è la chiusura sotto unione delle strutture dei predicati componenti. Formalmente, su un dominio X condiviso, il predicato disgiunto « P o Q » ha relazione di tolleranza $\sim_{\{PVQ\}}$ tale che $x \sim_{\{PVQ\}} y$ se $x \sim_P y$ oppure $x \sim_Q y$ (o entrambi). Il grafo di tolleranza $G(X, \sim_{\{PVQ\}})$ è l'unione dei grafi $G(X, \sim_P)$ e $G(X, \sim_Q)$. Il numero di cicli indipendenti del grafo unione è al più la somma dei cicli dei due grafi (per la disuguaglianza del numero ciclomático sotto unione di grafi), dunque $\text{rk}(H_1(S(X, \sim_{\{PVQ\}}))) \leq \text{rk}(H_1(S(X, \sim_P))) + \text{rk}(H_1(S(X, \sim_Q)))$. La vaghezza della disgiunzione è al più la somma delle vaghezze dei componenti — in linea con l'intuizione che un predicato disgiunto («è bello o buono») non sia più difficile da classificare di quanto lo siano i due predicati componenti presi insieme.

Il numero di risoluzione della disgiunzione soddisfa invece la disuguaglianza opposta: $\varepsilon^*(\sim_{\{PVQ\}}) \leq \min(\varepsilon^*(\sim_P), \varepsilon^*(\sim_Q))$, perché il cammino monotono minimo nel DAG unione è al massimo uguale al cammino monotono minimo in ciascuno dei DAG componenti. In termini linguistici, la disgiunzione di due predicati vaghi è (debolmente) meno vaga di ciascuno dei predicati componenti: «è alto o giovane» è più facile da soddisfare di «è alto» o di «è giovane» separatamente, e questa facilità si traduce in un numero di risoluzione minore.

10 Vaghezza e apprendimento automatico del linguaggio

Nell'ultimo decennio, la linguistica computazionale è stata profondamente trasformata dai modelli linguistici di grandi dimensioni (Large Language Models, LLM) basati sull'architettura transformer. Questi modelli, addestrati su enormi corpora di testo umano, apprendono rappresentazioni vettoriali (embeddings) dei token linguistici e producono testo in modo statisticamente coerente con i dati di addestramento. La domanda naturale, alla luce della teoria sviluppata nel presente saggio, è: come codificano questi modelli la vaghezza linguistica, e in che misura le strutture di tolleranza indotte sono coerenti con quelle documentate dalla psicofisica?

Un'osservazione fondamentale è che i modelli linguistici operano su spazi vettoriali continui di alta dimensione (tipicamente, spazi $\mathbb{R}^d$ con $d = 768, 1024, 4096$ o più). La rappresentazione di un termine linguistico come vettore in $\mathbb{R}^d$ implica una metrica

(tipicamente, la cosine similarity o la distanza euclidea) che induce una struttura di tolleranza sullo spazio degli embedding. La relazione di tolleranza tra due parole w_1 e w_2 nei loro embedding $e_1, e_2 \in \mathbb{R}^d$ è definita naturalmente come:

$$w_1 \sim_r w_2 \iff \|e_1 - e_2\|_2 \leq r$$

per una soglia r determinata dal contesto applicativo. Il grafo di tolleranza $G(W, \sim_r)$ sullo spazio delle parole W ha come vertici i termini linguistici e come archi le coppie a distanza $\leq r$ nello spazio degli embedding. La filtrazione di Vietoris-Rips su questo grafo produce un sorites barcode che codifica la struttura semantica dello spazio vettoriale.

La connessione con la teoria della vaghezza è la seguente. I termini che denotano predicati vaghi con dominio ordinato ("alto", "caldo", "vecchio") dovrebbero, se il modello ha appreso la struttura semantica del linguaggio naturale, avere strutture di tolleranza (nello spazio degli embedding) che riflettono la struttura di tolleranza psicofisica del predicato. In particolare, ci si attende che il sorites barcode dello spazio degli embedding, ristretto al sottospazio semantico rilevante, abbia una barra principale lunga in PD_0 — corrispondente alla connessione topologica del continuum di significati del predicato — e un numero di risoluzione ε^* inversamente proporzionale alla JND del predicato.

Questa previsione è empiricamente verificabile. Si può costruire lo spazio di tolleranza $(X_P, \sim_r^{\{\text{emb}\}})$ sulle istanze di un predicato P (per esempio, le 1000 frasi più frequenti che contengono «alto») usando come distanza la distanza coseno nello spazio degli embedding contestuali (per esempio, gli embedding di BERT o di GPT-4). Il sorites barcode di questo spazio può essere calcolato con gli algoritmi di omologia persistente (librerie come Ripser, Gudhi, o Giotto-TDA). Il confronto tra il barcode empirico e quello teorico (derivato dalla struttura di JND psicofisica del predicato) permette di misurare quanto fedelmente il modello linguistico ha codificato la vaghezza strutturale del predicato.

Un risultato teorico di rilievo in questa direzione riguarda la dimensionalità intrinseca dello spazio degli embedding dei predicati vaghi. La Proposizione 16 del Capitolo 7 mostra che per una griglia d -dimensionale di lato n , $\text{rk}(H_1)$ cresce come $d \cdot n^{d-1}$ per n grande. Se i modelli linguistici codificano correttamente la struttura multidimensionale dei predicati valutativi, il sorites barcode degli embedding di questi predicati dovrebbe avere un numero di cicli indipendenti in H_1 consistente con la dimensionalità intrinseca dello spazio semantico del predicato. Questa è una previsione

concreta e quantitativa che la teoria degli spazi di tolleranza formula sul comportamento degli embeddings linguistici.

Un aspetto più sottile riguarda la stabilità. Il Teorema 12 del Capitolo 9 (stabilità del diagramma di persistenza) garantisce che piccole perturbazioni della metrica producono piccole perturbazioni del barcode. Nei modelli linguistici, la metrica nello spazio degli embedding varia tra modelli diversi (anche dello stesso tipo ma con pesi diversi per effetto dell'inizializzazione random) e tra contesti diversi (gli embedding contestuali dipendono dal contesto della frase, non solo dal token). La stabilità del sorites barcode implica che, se due modelli o due contesti producono metriche vicine, i loro barcodes della vaghezza sono vicini. Una divergenza grande dei barcodes indica una codificazione strutturalmente diversa della vaghezza — un segnale di inconsistenza semantica del modello rispetto al linguaggio umano.

11 Implicazioni per la pragmatica e la teoria del significato

La teoria matematica della vaghezza sviluppata nel presente saggio ha implicazioni non soltanto per la semantica formale (il significato delle espressioni linguistiche in relazione alla struttura del mondo) ma anche per la pragmatica (l'uso del linguaggio in contesti comunicativi specifici). Queste implicazioni riguardano in particolare la funzione comunicativa della vaghezza — il fenomeno per cui i parlanti usano deliberatamente espressioni vaghe per motivi strategici — e le condizioni di felicità delle asserzioni vaghe.

La vaghezza deliberata, studiata da Channell (1994), Hirst e Albright (2002), e più recentemente da van Deemter (2010), è il fenomeno per cui i parlanti usano termini vaghi anche quando avrebbero la possibilità di essere precisi. «Arriverò tra poco» invece di «arriverò tra 14 minuti»; «il libro costa qualcosa» invece di «il libro costa 23,50 euro». La funzione comunicativa di questo uso è spesso la riduzione dell'impegno informativo: una asserzione vaga è meno falsificabile di una precisa e quindi più sicura in condizioni di incertezza.

Nel quadro della teoria degli spazi di tolleranza, la vaghezza deliberata corrisponde alla scelta, da parte del parlante, di usare un predicato con ε^* più grande del necessario per la distinzione comunicativa rilevante. Se il parlante sa che l'oggetto di cui parla si trova nella zona di valori $f(x) \in [a, b] \subseteq [0, 1]$ con $b - a \ll 1$ (conosce il valore con precisione alta), ma usa un predicato con $\varepsilon^* \gg b - a$, la zona di confine del predicato usato è molto

più larga della «zona di incertezza» fattuale. Il risultato è un'asserzione informativamente debole ma pragmaticamente sicura.

Un aspetto opposto è la vaghezza forzata: situazioni in cui il parlante non ha altra scelta che usare un predicato vago perché non esiste, nel sistema linguistico a disposizione, un predicato con struttura di tolleranza sufficientemente fine per distinguere il caso in questione. In questi casi, il numero di risoluzione ε^* del predicato disponibile è una misura della «risoluzione semantica» del sistema linguistico rispetto al dominio di oggetti. Sistemi linguistici con lessici più ricchi per un dominio hanno predicati con ε^* più piccoli — e quindi maggiore capacità discriminativa — su quel dominio.

Le condizioni di felicità delle asserzioni vaghe, nella tradizione della teoria degli atti linguistici di Austin (1962) e Searle (1969), richiedono che il parlante sia in grado di sostenere sinceramente la verità dell'asserzione. Per le asserzioni di predicati vaghi nella zona di confine — «Gianni è alto» per un Gianni di 178 cm — la condizione di sincerità non può essere «il parlante crede che Gianni è alto sia vero», perché nella zona di confine il valore di verità non è determinato. La condizione di sincerità appropriata, alla luce della teoria sviluppata, è: il parlante deve credere che il valore della funzione di membership $f(\text{Gianni})$ sia nel range $\varepsilon^*(\sim_{\text{alto}})$ -compatibile con la verità del predicato nel contesto dato. Più precisamente, il parlante deve credere $f(\text{Gianni}) \geq 1 - \varepsilon^*(\sim_{\text{alto}})$, il che corrisponde all'affermazione che Gianni si trova nel range di valori per cui il predicato è « ε^* -vero».

Questa riformulazione della condizione di sincerità per le asserzioni vaghe ha il vantaggio di essere operativa — specifica quantitativamente quando un'asserzione vaga è «sincera» — e di essere compatibile con tutte e tre le grandi tradizioni filosofiche sulla vaghezza. Per l'epistemicismo, la condizione di sincerità si riduce alla condizione classica (il parlante crede che il valore sia sopra la soglia netta, senza saperlo esattamente), con la vaghezza che emerge dall'incertezza epistemica sulla posizione della soglia. Per la supervalutazione, la condizione di sincerità è che l'asserzione sia supervera in un insieme sufficientemente ampio di precisazioni. Per le logiche fuzzy, la condizione di sincerità è che il valore di membership $f(x)$ sia $\geq 1 - \varepsilon^*$, che è esattamente la formulazione qui proposta.

12 La vaghezza come risorsa: entropia e capacità comunicativa

Un ultimo aspetto della vaghezza linguistica, spesso trascurato nelle trattazioni filosofiche ma rilevante per la linguistica computazionale e la teoria dell'informazione, è la sua funzione come risorsa comunicativa. La vaghezza non è soltanto un «difetto» da eliminare o gestire, ma una caratteristica strutturale del linguaggio naturale che consente la comunicazione efficiente in condizioni di incertezza e variabilità.

Nella prospettiva della teoria dell'informazione di Shannon, ogni sistema di comunicazione si caratterizza per la sua capacità C : il massimo tasso di trasmissione di informazione per canale per uso del canale, misurato in bit. Per un predicato vago usato come canale di comunicazione, la capacità dipende dalla struttura di tolleranza: quante distinzioni il predicato può trasmettere in modo affidabile (con probabilità di errore $< \delta$), tenuto conto della struttura di indistinguibilità.

Il numero di risoluzione $\varepsilon^* = 1/L_mono_min$ fornisce una misura diretta della capacità di trasmissione del predicato. Un predicato con $L_mono_min = L$ può trasmettere in modo affidabile al più L distinzioni lungo il dominio ordinato: le $L+1$ regioni $[0, 1/L)$, $[1/L, 2/L)$, ..., $[(L-1)/L, 1]$ della scala di membership. La capacità del predicato come canale di comunicazione è:

$$C_P = \log_2(L_mono_min + 1) \quad \text{bit}$$

Per il predicato «è alto» con $L_mono_min = 85$ (calcolato nella Sezione 16.2), la capacità è $C_{\text{alto}} = \log_2(86) \approx 6,43$ bit. Questo significa che il predicato «è alto» può distinguere, in linea di principio, circa $2^{\{6,43\}} \approx 86$ livelli di altezza in modo comunicativamente affidabile — un valore che corrisponde al numero di passi del cammino monotono minimo, come atteso.

Per i predicati multidimensionali, la capacità cresce molto più rapidamente. Un predicato bidimensionale $m \times n$ ha cammino monotono minimo di lunghezza $m+n-2$, dunque capacità $\log_2(m+n-1)$ bit. Per $m = n = 10$, la capacità è $\log_2(19) \approx 4,25$ bit, che è inferiore alla capacità del predicato unidimensionale con 86 livelli. Il paradosso apparente — che un predicato bidimensionale su 100 elementi abbia capacità minore di un predicato unidimensionale su 86 elementi — si risolve osservando che la capacità qui calcolata è quella lungo una singola dimensione ordinata. La capacità totale del predicato bidimensionale, misurata rispetto all'intero spazio di oggetti (100 oggetti), è la somma delle capacità lungo ciascuna «direzione» del cammino ottimale, e dipende dalla struttura complessiva del DAG multidimensionale.

L'entropia della distribuzione di probabilità dei valori di membership — che misura l'incertezza residua del predicato applicato a istanze campionate casualmente dal dominio — è correlata ma distinta dalla capacità. Per la distribuzione uniforme sulle $L+1$ regioni di membership, l'entropia è $\log_2(L+1) = C_P$ bit, pari alla capacità. Per distribuzioni non uniformi (per esempio, se la maggior parte degli individui nel campione è nella zona di confine, come accade per l'altezza vicino alla media della distribuzione), l'entropia è minore della capacità, e la «utilità comunicativa effettiva» del predicato è ridotta rispetto al massimo teorico.

Questo quadro permette di collegare la teoria della vaghezza linguistica alla teoria statistica dell'apprendimento e al principio del minimo descrittivo lunghezza (Minimum Description Length, MDL) di Rissanen (1978). Il predicato con ε^* ottimale per un dominio dato è quello che minimizza la lunghezza della descrizione MDL degli elementi del dominio: né troppo vago (che non permette di distinguere elementi distinti) né troppo preciso (che richiede un lessico eccessivamente fine per codificare la variabilità del dominio). La vaghezza «naturale» dei predicati linguistici, documentata dalla JND psicofisica, emerge come la scelta che ottimizza il trade-off tra capacità comunicativa e costo cognitivo della distinzione.

13 Risultati principali e sintesi

Il quadro matematico sviluppato nei Capitoli 0–15 del presente saggio permette ora, nell'ottica di questa sezione conclusiva, di articolare la struttura della vaghezza linguistica in modo preciso e comprensivo. La vaghezza linguistica è:

1. Una proprietà strutturale dei predicati, non un difetto del linguaggio né un fenomeno di ignoranza. Il complesso di Sorite $S(X, \sim_P)$ è l'oggetto matematico che codifica questa proprietà strutturale, e le sue proprietà topologiche — connessione, cicli, omologia — determinano la struttura della vaghezza.

2. Misurabile quantitativamente attraverso il numero di risoluzione $\varepsilon^*(\sim_P) = 1/L_mono_min$, che è la minima incoerenza graduata da accettare per un assegnamento con contrasto. Questo numero è calcolabile in tempo $O(|X| + |E|)$ sull'insieme degli elementi del dominio e sulla relazione di tolleranza.

3. Stratificabile in livelli attraverso il sorites barcode, che descrive la struttura multiscala della vaghezza: le barre lunghe corrispondono a distinzioni semantiche stabili e

robuste, quelle brevi a distinzioni fragili dipendenti dal contesto. L'entropia del barcode quantifica il grado di vaghezza di ordine superiore.

4. Composizionale in modo non banale: la vaghezza di predicati composti non è la semplice somma delle vaghezze dei predicati componenti. La composizione crea nuova topologia (nuovi cicli in H_1) che non era presente nei predicati elementari, come dimostrato dalla teoria delle griglie multidimensionali (Teorema 6 e Proposizione 16).

5. Stabile rispetto a perturbazioni contestuali: il Teorema 12 garantisce che piccole variazioni della struttura metrica del dominio producono piccole variazioni del sorites barcode, misurate dalla distanza bottleneck. Questo spiega la robustezza della comunicazione linguistica con predicati vaghi nonostante la variabilità contestuale.

6. Connessa alla struttura spettrale del grafo di tolleranza attraverso il gap spettrale μ_1 del laplaciano normalizzato. Un gap spettrale piccolo corrisponde a una zona di confine pronunciata (collo di bottiglia nella struttura del grafo), mentre un gap grande corrisponde a un confine relativamente netto. La disuguaglianza di Cheeger ($h^2/2 \leq \mu_1 \leq 2h$) quantifica questo legame.

Il risultato principale di questa sezione — che può essere enunciato come un teorema meta-linguistico — è che la vaghezza linguistica è necessaria in qualunque sistema di predicati su domini percettivamente continui, cioè su domini su cui opera la struttura JND della percezione. Formalmente: per qualunque dominio (X, d) con metrica non discreta e qualunque soglia $r > 0$, il grafo di tolleranza $G(X, \sim_r)$ con $\sim_r$ uguale alla tolleranza metrica di raggio r è connesso (per X connesso nella topologia indotta da d), dunque $\varepsilon^*(\sim_r) > 0$ per ogni $r > 0$ e il paradosso di Sorite si applica a qualunque predicato con soglia di indistinguibilità positiva su X . La vaghezza non è accidentale ma strutturale nella relazione tra il linguaggio e la percezione continua del mondo.

Teorema meta-linguistico (vaghezza strutturale). *Sia X un insieme finito di istanze percettive di un predicato P , dotato di una metrica $d: X \times X \rightarrow \mathbb{R}_{\geq 0}$ tale che il grafo di adiacenza $G(X, \sim_r)$ sia connesso per ogni $r > 0$ abbastanza piccolo (condizione soddisfatta per qualunque X denso in un dominio connesso). Per ogni soglia di indistinguibilità $r > 0$ fissata dalla struttura psicofisica della percezione, il predicato P è vago in senso strutturale: non esiste nessun assegnamento bivalente coerente con contrasto. Il grado di vaghezza $\varepsilon^*(\sim_r) = 1/L_{\text{mono_min}}(\sim_r)$ è strettamente positivo e tende a 0 al tendere di r a 0, ma è positivo per qualunque $r > 0$ fissato. La vaghezza è*

dunque ineliminabile per qualunque predicato su domini percettivamente continui con soglia di discriminazione positiva: essa non è un difetto del linguaggio ma una conseguenza strutturale della relazione tra i predicati linguistici e la struttura della percezione.

Il teorema meta-linguistico non nega la possibilità di costruire sistemi linguistici artificiali privi di vaghezza: basta scegliere $r = 0$ (discriminazione infinita), il che corrisponde a predicati definiti intensionalmente tramite condizioni necessarie e sufficienti esatte (definizioni stipulative). Ma tali sistemi operano al di fuori della struttura psicofisica della percezione umana, e i loro predicati non sono acquistabili per ostensione: non possono essere appresi mostrando esempi, ma soltanto tramite definizioni esplicite. Il linguaggio naturale, che si apprende per ostensione, non può essere privo di vaghezza strutturale per la stessa ragione per cui la percezione non può essere priva di JND.

In conclusione, la teoria combinatoria e topologica della vaghezza sviluppata in questo saggio offre una risposta precisa alla domanda che aveva aperto il percorso: cosa è la vaghezza linguistica, matematicamente? È la topologia non banale del complesso di Sorite associato al predicato — una topologia che emerge necessariamente dalla struttura dell'indistinguibilità percettiva degli oggetti del mondo, che è misurabile attraverso il numero di risoluzione ϵ^* e il sorites barcode, e che è stabile, compositazionale, e connessa alla struttura spettrale del grafo di tolleranza. Non è la vaghezza un'imperfezione del pensiero né una trappola del linguaggio: è la traccia matematica che la continuità del mondo lascia sulla struttura discreta dei predicati con cui lo categorizziamo.

Riferimenti bibliografici aggiuntivi

- Austin, J. L. (1962). *How to Do Things with Words*. Oxford University Press. Oxford.
- Barker, C. (2002). The dynamics of vagueness. *Linguistics and Philosophy*, 25(1), 1–36.
- Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (a cura di), *Concepts and Conceptual Development*. Cambridge University Press. Cambridge.

- Channell, J. (1994). *Vague Language*. Oxford University Press. Oxford.
- Cohen-Steiner, D., Edelsbrunner, H., e Harer, J. (2007). Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1), 103–120.
- Cresswell, M. J. (1977). The semantics of degree. In B. Partee (a cura di), *Montague Grammar*. Academic Press. New York.
- Edelsbrunner, H., Letscher, D., e Zomorodian, A. (2002). Topological persistence and simplification. *Discrete & Computational Geometry*, 28(4), 511–533.
- Edgington, D. (1997). Vagueness by degrees. In R. Keefe e P. Smith (a cura di), *Vagueness: A Reader*. MIT Press. Cambridge, MA.
- Fara, D. G. (2000). Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics*, 28(1), 45–81.
- Faye, J.-P. (1972). *Langages totalitaires*. Hermann. Paris.
- Fine, K. (1975). Vagueness, truth and logic. *Synthese*, 30(3–4), 265–300.
- Gal, S. R. (2005). Real root conjecture fails for five- and higher-dimensional spheres. *Discrete & Computational Geometry*, 34(2), 269–284.
- Gardner, H. (1983). *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books. New York.
- Garey, M. R. e Johnson, D. S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman. San Francisco.
- Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 18(4), 441–461.
- Hatcher, A. (2002). *Algebraic Topology*. Cambridge University Press. Cambridge.
- Hirst, G. e Albright, A. (2002). Approximate and logical reasoning in natural language processing. In D. Hirst e G. McRae (a cura di), *Approximation in Language and Thought*. John Benjamins. Amsterdam.
- Kennedy, C. (1999). *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. Garland. New York.

- Kennedy, C. e McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, 81(2), 345–381.
- Latschev, J. (2001). Vietoris-Rips complexes of metric spaces near a closed Riemannian manifold. *Archiv der Mathematik*, 77(6), 522–528.
- Leray, J. (1945). Sur la forme des espaces topologiques et sur les points fixes des représentations. *Journal de Mathématiques Pures et Appliquées*, 24, 95–167.
- Levin, D. A., Peres, Y., e Wilmer, E. L. (2009). *Markov Chains and Mixing Times*. American Mathematical Society. Providence.
- Łukasiewicz, J. (1920). O logice trójwartościowej. *Ruch Filozoficzny*, 5, 170–171.
- Mehlberg, H. (1958). *The Reach of Science*. University of Toronto Press. Toronto.
- Montague, R. (1973). The proper treatment of quantification in ordinary English. In J. Hintikka, J. Moravcsik e P. Suppes (a cura di), *Approaches to Natural Language*. Reidel. Dordrecht.
- Munkres, J. R. (1984). *Elements of Algebraic Topology*. Addison-Wesley. Menlo Park.
- Niyogi, P., Smale, S., e Weinberger, S. (2008). Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39(1–3), 419–441.
- Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5), 465–471.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104(3), 192–233.
- Rothschild, D. e Segal, G. (2009). Indexical predicates. *Mind & Language*, 24(4), 467–493.
- Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press. Cambridge.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423.
- Sorensen, R. (1988). *Blindspots*. Clarendon Press. Oxford.

Sternberg, R. J. (1985). *Beyond IQ: A Triarchic Theory of Human Intelligence*. Cambridge University Press. Cambridge.

van Deemter, K. (2010). *Not Exactly: In Praise of Vagueness*. Oxford University Press. Oxford.

van Fraassen, B. (1966). Singular terms, truth-value gaps, and free logic. *Journal of Philosophy*, 63(17), 481–495.

Williamson, T. (1994). *Vagueness*. Routledge. London.

Wittgenstein, L. (1953). *Philosophische Untersuchungen / Philosophical Investigations*. Blackwell. Oxford.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353.

Zomorodian, A. e Carlsson, G. (2005). Computing persistent homology. *Discrete & Computational Geometry*, 33(2), 249–274.